THE DIFFUSION OF NEW PROCESS INNOVATIONS IN U.K. MANUFACTURING INDUSTRIES.

by

Stephen Davies

submitted for the degree of Ph. D. in Economics

at

The University of Warwick.

September 1975.

Summary.

This thesis comprises a theoretical and empirical study of the spread of new techniques within industry (the so-called diffusion process.) The major aim has been to provide an economic explanation for differences between industries in the speed with which they adopt new techniques or innovations.

The theoretical underpinning of the study is a general model of the diffusion process based on an explicit theory of individual firms' decision making behaviour in this context. This model is built around the discussions of the early chapters on the technological characteristics of new process innovations and the results of past work on the nature of the adoption decision at the firm level.

The empirical part of the thesis is concerned with testing the various predictions generated by the model using data collected on the diffusion of 22 major innovations in various U.K. industries since the war. A number of hypotheses appear to be confirmed: at the individual firm level, behaviour is partly determined by the firm's size; at the industry level, competitive structure and aggregate demand conditions appear to influence the speed of diffusion, and further, the characteristics of the innovations themselves affect not only the speed of diffusion but also the shape of the diffusion growth curve.

The thesis constitutes the first large scale empirical study of diffusion in the U.K. industry. Theoretically, it provides the first model of diffusion to be based on economic decision making rather than the mechanistic models of epidemics which have been used previously in this area.

CONTENTS

Contents continued...

ACKNOWLEDGEMENTS

Whilst writing this thesis I have benefited from the encouragement and assistance of a number of people. My sincere thanks are due to my supervisor, Graham Pyatt: his guidance and support have been much appreciated.

The onerous task of data collection has been substantially eased by the kind co-operation of many people in various industries. In this connection, I am particularly indebted to David Worswick, the Director of the National Institute of Economic and Social Research, who generously made available to me large quantities of data on the diffusion of a number of innovations.

Finally, I should like to thank Elizabeth Holloway for typing the final draft and my wife for typing earlier drafts and for her patience and encouragement throughout the duration of my research.

List of Tables

List of Figures

Chapter 1 : Introduction.

Edwin Mansfield claimed in 1968 that "After many years of neglect, technological change is receiving the attention that it deserves"[1], and Kennedy and Thirlwall's summary article of 1972 [2] confirms that this attention has been sustained: a head count of the number of books and articles on 'technical progress' shows a continued exponential growth.

It is not difficult to identify the raison d'etre for this surge of academic effort. There had been a tendency for much of this century, at least, to assume technical progress to be God-given and thus exogeneous to the system, no matter how broadly that system was defined. Whilst this exogeneity simplified the analysis of economic growth and thus welfare, it also reduced the feasible area for policy implications. With the flurry of empirical research of the late nineteen-fifties [3], purporting to show that 'technical progress' accounted for at least 80% of the growth in labour productivity in the U.S., something of a corporate decision can be seen to have followed from the Economics profession. If, indeed, technical progress was so important, then it was no longer satisfactory to assume it to be like 'Manna from Heaven' - rather, it must be explained. In broad terms then, this upsurge in interest noted by Mansfield can be seen as an attempt to 'endogenize technical progress'.

The present thesis follows this trend. It takes, as its subject, one particular aspect of technical progress, namely the diffusion of process innovations, and attempts a predominantly empirical explanation. The 1972 survey article mentioned above reports a very strong emphasis on empirical

1. E. Mansfield, "Industrial Research and Technological Innovation."Norton, New York, 1968 (Preface.)

2. C. Kennedy and A. Thirlwall, "Surveysin Applied Economics:Technical Progress". Economic Journal, March 1972.

3. Most notably, R. Solow, "Technical Change and the Aggregate Production Function", Review of Economics and Statistics, August 1957. But also: M. Abramowitz, "Resource and output Trends in the U.S. since 1870", American Economic Association, Papers, May 1956, and S.Fabricant,"Basic Facts on Productivity Change," N.B.E.R. Columbia University Press, New York, 1959.

analysis. Indeed, the charge is often made that this emphasis has resulted in too much ad-hoc theorizing. Consequently, a large proportion of this study will be devoted to developing a fully worked theory of diffusion before any empirical work is attempted.

1. Some points of definition.

Diffusion of a new innovation represents the final link in what might be termed the chain of technical change. First an invention is made, either on the basis of new scientific knowledge, or, more commonly, using well known scientific principles. The innovation stage is reached once this invention has been modified, as necessary, and introduced commercially for the first time by one firm, often called the innovator. Diffusion takes place as other firms, in the same industry, imitate[1] and adopt the innovation themselves. A study of the diffusion process is, then, a study of the spread of a new technique or product within an industry, once the innovation has been introduced by one firm in that industry. Generally, little will be said here about the invention or innovation stages.

In this study only process innovations will be considered. To follow Blaug's amusing, but non-rigorous, definition, a process innovation is 'a novel way of making old goods' whilst a product innovation involves 'old ways of making novelties'.[2] Generally, a process innovation involves a reduction in costs per unit of output, despite the fact that input prices remain unchanged. Having said this, a new process sometimes implies, in practice, a change in the nature of the end product for the potential adopter.[3] Furthermore, if the innovation is made available to the potential adopter by a capital goods sector, then for the latter, the innovation is a product innovation. For instance, a new type of steel furnace, is a process for the Steel industry but a new product for the furnace-making industry.

1. J. Schumpeter was, perhaps, first to use this three way classification of invention, innovation and diffusion see, "The theory of economic development", Harvard University Press and "Capitalism, socialism and democracy" Harper and Row, New York 1942.

2. M. Blaug "A Survey of the Theory of Process Innovations", Economica, Feb.1963.

3. For examples of this phenomenon, see chapter 4. p.26.

Although the word diffusion will be used freely and frequently in this thesis, nearly always it is the imitation process that will be considered. There are three dimensions to the diffusion process: imitation or inter-firm diffusion refers to the spread of the process from firm to firm; intra-firm diffusion to the spread of processes within individual firms; and overall diffusion to the spread throughout the industry, as a whole, encompassing both inter- and intra-firm diffusion. Thus, inter-firm diffusion is measured by the proportion of firms in the industry that have adopted, intra-firm diffusion by the proportion of any one firm's output produced using the new process, and overall diffusion by the proportion of the total industry output that is produced using the new process. From the point of view of the change in industry productivity resulting from the spread of the new innovation, overall diffusion is, of course, the most pertinent measure. However, as just mentioned, it is the imitation process which is considered almost exclusively in this work. This choice can be justified on three grounds. Firstly, it has proved very difficult to collect data on inter-firm diffusion, yet this only requires knowledge of the date of first adoption of the new technique for each firm in the relevant industry. It seems likely that the more detailed data needed to study intra-firm and overall diffusion (i.e. data on the spread of the technique within those firms) would be more difficult to collect. Almost certainly, most firms would be either incapable or unwilling to provide such data; quite definitely, there is no published data available. Secondly, intra-firm diffusion is often trivial: for many innovations, if a firm adopts at all, then it must adopt 100%. For instance, a lumpy innovation, such as computer typesetting, requires only one computer to be installed for the firm to produce all of its output using the new process.[1] In such cases, inter-firm diffusion is equivalent to overall diffusion (given a knowledge of firms'

1. As opposed to, say, shuttleless looms, for which the initial adoption decision may entail only 10 conventional looms being replaced, leaving perhaps 90 to be replaced at later dates.

sizes.) Thirdly, it is the initial decision to adopt an innovation that is particularly worthy of analysis, involving, as it does, a study of decision-making under uncertainty. Once a firm has made the decision to adopt, its later decisions, on the speed at which to switch the rest of its output to the new process, involve a different set of considerations requiring a separate analysis.

To summarise briefly, the subject matter of the thesis is the imitation process, hereafter referred to as (inter-firm) diffusion. This defines the spread of a new process amongst firms within any one industry and takes no account of the scale of their adoption. The initial introduction of the process, by the innovating firm, is taken as given and is the starting point of the diffusion process.

2. The Methodology and Data.

The broad methodology is borrowed from Griliches' and Mansfield's work. An empirical examination of the determinants of the diffusion rate is undertaken using cross-section analysis, in which each observation relates to the diffusion of a separate innovation in (where possible) a different industry. In order to generate these observations, time series data are required on the diffusion of each of these innovations. The empirical work thus comprises two stages: for each innovation, various standard growth curves (Mansfield used the logistic) are fitted to the time series for diffusion. The estimated parameters of these curves are then used in a second stage as the dependent variable, (representing the rate of diffusion), in cross-section regressions, in which the independent variables are various characteristics of the innovations and industries involved. It is these variables which are considered to be the determinants of the rate of diffusion at the micro or industry level.

Clearly, for the explanation to be at all general, data is needed on a large number of innovations diffusing in a wide range of industries. Consequently, data requirements have been high and a general aim has been to collect sufficient information on as many innovations as possible.

In the event, it has proved possible to acquire data on 22 innovations, which is probably a large enough number to permit the required level of generalisation.

A first data requirement for each innovation is knowledge of the number of firms having adopted at yearly intervals. In addition, in order to build a realistic theoretical model, a fairly deep understanding is required of the technical aspects of these innovations. This necessitates the collection of a large amount of technical and scientific information, both quantitative and qualitative. Third, a fair amount of information is needed about the industries concerned. Of course, much of this industry-level information is readily available in easily accessible sources. However, certain problems arise because the relevant industries are of a highly disaggregated nature e.g. the malting, clay brickmaking and provincial newspaper industries. As the model developed requires detailed information on, for instance, the size distributions of firms, which are not published at this level of disaggregation, this also presents a substantial data collection exercise.[1]

The main practical problems, however, concern the diffusion and technical data. An early pilot study illustrated quite clearly that many firms are loathe to disclose even the minimum necessary information, i.e. the date of their adoption of a new process. Alternative sources of data were obviously required. For a small number of important innovations, certain scientific journals provide data on the spread of the innovation and carry adequate technical descriptions; for a few others, trade and research associations provide sufficient information, but this still leaves the sample short of the desired number of innovations. However, a fair amount of useful data has already been collected by other economists working in this field. The National Institute of Economic and Social Research, T. Scott and J. Metcalfe

1. Relying heavily on trade associations and trade journals.

have each studied the diffusion of certain innovations in case studies (N.I.E.S.R. have, in fact, studied a dozen innovations, as their contribution to an international study of diffusion in six countries.[1]) Their generous co-operation in making this data available to me has therefore brought the sample size up to twenty two innovations in thirteen industries. As such, a level of generalisation far in excess of that achieved in past studies should now be possible.

Because of this generous help, data problems have been contained at a manageable level. Nevertheless, a sizeable residual data collection has been necessary, particularly on the technical aspects of the individual innovations. The main sources used have been trade and scientific journals and correspondence with a large number of machine makers (responsible for marketing the innovations.) The last mentioned have been particularly helpful and patient.

A more detailed discussion of data collection and a presentation of the data appears in Appendices 1 - 5. Meanwhile a list of the innovations in the sample is set out in table 1.1.

1. See Appendix one. It should be stressed that although a substantial portion of the data has been 'borrowed', there is no question of a duplication of past research since each concentrated on the case-study approach, with the emphasis on in-depth descriptions of the particular innovations.

Table 1.1. The Sample Innovations.

Innovation	Abbreviation used	Industry in which diffusion occurred
Special presses for paper machines	S.P.	Paper and Board
Foils for paper machines	F.	" "
Synthetic fabrics for paper machines	S.F.	" "
Wet Suction boxes for paper machines	W.S.B.	" "
Process control by computer of paper-making.	P.C.B.C.	" "
Gibberellic Acid additive to barley	G.A.	Malting
Computer Typesetting	C.T.	Printing and publishing of provincial evening newspapers
Shuttleless looms	S.L.	Textile weaving
Electric hygrometer for sizing	E.H.	Lancashire textile weaving
Accelerated drying hoods for sizing	A.D.H.	" " "
Automatic size boxes for sizing	A.S.B.	" " "
Tufted carpet machines	T.C.	Carpet manufacture
Automatic track lines	A.T.L.	Car manufacture
New methods of steel plate cutting	S.P.C.	Shipbuilding
Numerically controlled machine tools	N.C.P.P.	Printing press manufacture
Numerically controlled machine tools	N.C.TURN.	Turning machine manufacture
Numerically controlled machine tools	N.C.TURB.	Turbine manufacture
Tunnel kilns	T.K.	Claybrick Manufacture
Basic oxygen process in steelmaking	B.O.P.	Iron and Steel
Continuous casting in steelmaking	C.C.	" "
Vacuum degassing in steelmaking	V.D.	" "
Vacuum melting in special steelmaking	V.M.	Special steels

3. Theoretical Requirement

Given the nature of the data and the empirical aims set out above, certain things are required of the theoretical model to be constructed. Its basic assumptions about the technical nature of the new process innovations must be appropriate to the technological data just mentioned; it must be capable of yielding predictions as to the shape of the typical diffusion growth curve; and it must suggest which variables should affect the parameters of the growth curve. Hopefully these variables will have policy implications, particularly with respect to competition policy and demand management. For instance, the sort of questions which should be answerable are 'what influences on the speed of diffusion are exerted by industrial structure, firm size and demand conditions?'[1]

4. A general outline

Chapter 2 - a summary of past research in this field - fulfills two main functions. As a backcloth to the thesis, it summarizes what is already known about the diffusion of new industrial processes. An assessment is also provided of the theoretical models available, which might be used in this study. The main empirical findings to date are that, typically, the diffusion growth curve follows an upward sloping S curve and differences between innovations in the slope of the curve can be, at least, partly explained by differences in the profitability of those innovations. Moreover, there is substantial evidence (often based on rather suspect statistical assumptions) that large firms, on average, adopt more quickly than small firms. To date, however, little is known of the influence of industrial structure. Two main criticisms apply to the theoretical models used, particularly the well-known epidemic model of Mansfield and Griliché́s. The economic content is often minimal, having nothing to say on the adoption decision at the firm level, and often, but not always, the discussion of technical factors is very limited. Two candidates are seen as potentially forming the basis of an

1. That is, demand for the product of the industry which is adopting the new process.

acceptable model with which to analyse the sample innovations: the simple vintage model of Salter and the probit model which has been used in the analysis of the diffusion of consumer durables.

Chapter 3 considers the technological characteristics of the sample innovations in order to establish whether the typical new process may be described by a number of stylized facts. In addition to summarizing the evidence presented in the technical appendix 1, it draws on a number of past research findings on the nature of new technology. A number of important conclusions emerge. First, it seems that the sample is fairly broad-based, encompassing a wide range of different industries and innovations. It does appear that these innovations tend to have certain common characteristics; these suggest that the epidemic model of diffusion may be inappropriate on technological grounds. Moreover, the technological assumptions of the simple vintage model also do not seem to be widely applicable.

Chapter 4 analyzes the adoption decision at the firm level. On the basis of past research findings on a) the methods of investment evaluation actually used by firms and b) the diffusion of information within industries, it is suggested that a behavioural model of decision making may be most relevant in this context. The influence of industrial structure is considered on a theoretical level, as is the potential role of those firms which supply the new innovation.

Chapter 5 contains a mathematical statement of the model which is based on the findings of the two previous chapters. This model may be seen as an extension of the probit model mentioned above. Three types of prediction emerge which may be tested against the data collected.

The growth curve of diffusion may take one of two basic forms - cumulative normal or cumulative lognormal - depending on the type of the innovation. Further, superimposed on these underlying forms, there may be cyclical fluctuations due to the influence of the trade cycle. Chapter 6 tests these predictions for the time series data collected, and it is concluded that the model provides an encouraging explanation which is superior to that produced by the epidemic model. The data used in this chapter is presented and discussed in Appendices 2 and 3.

A second prediction of the model is that, for any innovation at all points in time, the probability of adoption should be positively related to firm size. Chapter 7 provides a test of this prediction and, again, the results are favourable. In this case, however, the limited quantity of data available prevents any rigorous testing. This data is presented and discussed in Appendix 4.

In chapter 8, the cross-industry predictions of the model are tested. It is found that much of the variance in the speed of diffusion across industries and innovations may be accounted for by differences in industrial structure and, as found by Mansfield and Griliches, the profitability of innovations. The regressions of this chapter require estimates of the parameters of the firm size distributions for each industry. These are calculated in Appendix 5. It appears that in all but one case, the observed distributions may be adequately described by the lognormal distribution. This is a particularly important finding as this assumption is necessary in the construction of the model in chapter 5. In Appendix 6 a discussion is presented as to the most appropriate way of measuring industrial structure.

In chapter 9 the implications of the results of the previous chapter are considered, with special reference to Government competition policy. Chapter 10 provides a summary of the main findings, an overall evaluation of the empirical success of the model and some suggestions for the direction of future research.

Chapter 2: A critical survey of past research.

This chapter serves two purposes. First, it attempts to summarise what is known already about the diffusion of new industrial processes. Second, it serves as a starting point in the search for a theoretical framework with which to analyse the diffusion of the innovations in the present sample. Consequently, so far as is possible, discussion of empirical findings will be separate from the analysis of the theory used in past work, although, obviously, empirics can rarely be discussed in isolation from the theory on which they are based.

Sections 2 to 4 survey what might be termed the mainstream literature, in which the main aim has been to analyse various aspects of diffusion performance using cross-section data. This may be thought of as incorporating three approaches. Most well-known is the <u>inter-industry/innovation approach</u> pioneered by Mansfield and Griliches: this amounts to studying the diffusion of one or more innovations in a number of industries and attempting to explain, empirically, the variance of the speed of diffusion in terms of differences in the attributes of the industries and innovations concerned. Alternatively, <u>the inter-firm approach</u>, also pioneered by Mansfield, concentrates on individual innovations diffusing in single industries and attempts to explain differences between firms in the time taken to adopt. In this case, firm-level characteristics are the explanatory variables. Third, the <u>international approach</u> attempts to explain international differences in the speed of diffusion of innovations in terms of the characteristics of the countries and industries concerned.

Section 1 acts as something of a preface to this discussion. A theoretical basis for much past research has been the assertion that the diffusion of new process innovations is analogous to the spread of infectious diseases. Therefore, as a backcloth to sections 2 to 4, section 1 provides a brief exposition of the most common mathematical model of epidemics, leading on to a discussion of the logistic curve which has been used so extensively in this area.

Section 5 considers the use of <u>stock-adjustment</u> models for the analysis of diffusion. They have been used not so much for comparative studies but more to analyse the time-path of diffusion for individual innovations.

A conclusion that emerges from the first five sections is that no totally acceptable[1] theoretical model exists in the literature on the diffusion of new processes. Section 6 broadens the scope of the survey to include a theoretical model that has been used in work on the <u>diffusion of new consumer durables</u>. It is considered that this model does form a potential basis for the analysis to be attempted in later chapters.

Finally, section 7 outlines an alternative embryonic model, using the <u>conventional vintage model</u> as propounded by Salter. It is noted, however, that the technological assumptions of this model are fairly crucial and that such an analysis would only be worthwhile here if these assumptions are valid for most of the innovations in the sample.

<u>1. A mathematical theory of epidemics and the logistic curve.</u>

The study of diffusion is not peculiar to economics amongst the social sciences. For instance, the spread of rumours, the use of new drugs, new teaching methods and steel axes by aborigine tribes have all been the subject of research by sociologists, medical sociologists, educationalists and anthropologists respectively.[2] One striking similarity exhibited by much of this research is the analogy often drawn to the spread of diseases. Consequently, a theoretical tool often used is one of the mathematical theories of epidemics. As reference to these theories is also consistently made by economists working in this area, an exposition of the simplest model of epidemics will provide a useful backcloth to the ensuing discussion in this chapter

1. Given the aims of this study, namely, an investigation of the role of certain industry-level variables such as market structure. Moreover, none of the research covered up to this point has much to say about the central issue of firm decision-making under uncertainty.

2. See E. Rogers, "Diffusion of innovations," Free Press of Glencoe, New York 1962.

particularly and, more generally, in the thesis as a whole.[1]

If at time t, m_t individuals in a population of n have contracted an infectious disease, then the number of individuals contracting it between times t and t+1 is proportionate to the product of the number uninfected and the proportion infected, both at time t.

$$m_{t+1} - m_t = \beta (n - m_t) \; m_t/n \qquad \beta > 0 \qquad (2.1.1.)$$

Thus the proportion of uninfected individuals who contract the disease in the time period is determined by the intensity of the infectiousness of the infecteds. This, in turn, may be defined as the product of the proportion already infected and the (constant) propensity of each non-infected to catch the disease (β).[2] Assuming t to t+1 is a very short time period, this may be stated as:

$$\frac{dm_t}{dt} \frac{1}{n - m_t} = \beta \frac{m_t}{n} \qquad (2.1.2.)$$

This differential equation has the solution:

$$\frac{m_t}{n} = \frac{1}{(1 + e^{-(\alpha + \beta t)})} \qquad (2.1.3.)$$

where α is a constant of integration.

This is the equation of the well-known logistic time curve (see figure) which has the following properties:

$$\lim_{t \to -\infty} (m_t/n) = 0 \qquad (2.1.3a.)$$

$$\lim_{t \to +\infty} (m_t/n) = 1 \qquad (2.1.3b.)$$

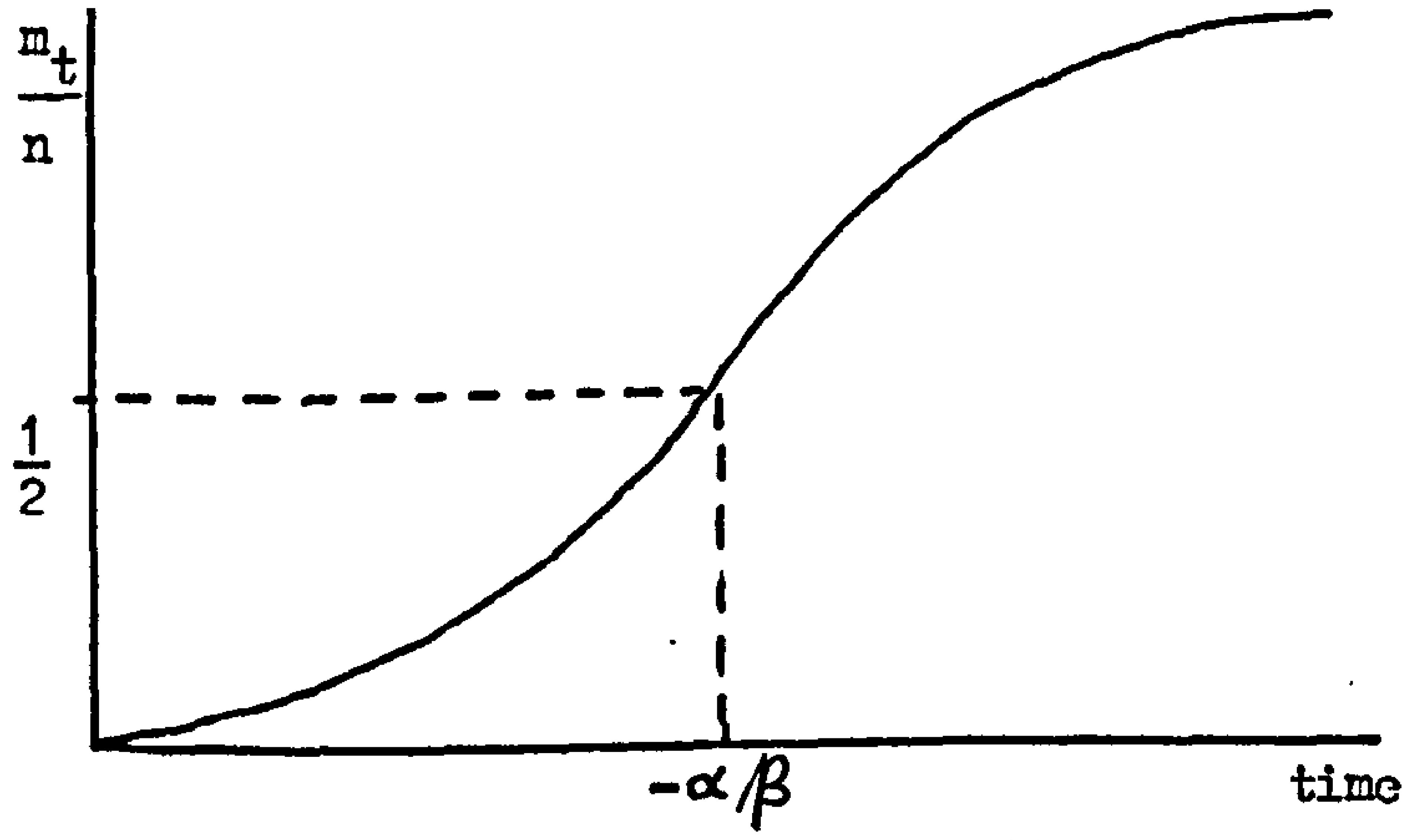
At $t = -\alpha/\beta$, $m/n = 1/2$, $d^2(m/n)/\, dt^2 = 0$, $d^3(m/n)/\, dt^3 \neq 0$ (2.1.3c.)

That is, the curve is symmetrical about its point of inflexion at $t = -\alpha/\beta$, $m/n = 1/2$.

1. This exposition follows, loosely and with amended notation, N. Bailey, 'Mathematical theory of epidemics,' Griffin London, 1957 and D.J. Bartholomew, 'Stochastic models for social processes.' Wiley, London 1970.

2. The magnitude of β will depend on the infectiousness of the disease and the frequency of social intercourse, amongst other things.

Figure 2.1.1. The logistic curve.



Usually, n is imposed in estimation and α and β may be estimated using weighted least squares[1] on the transformation:

$$\log \left\{ m_t / (n - m_t) \right\} = \alpha + \beta t \qquad (2.1.4.)$$

α is usually considered the lesser of the two parameters, only serving to locate the curve on the horizontal axis, whilst β is often defined as <u>the</u> rate of diffusion. This should not be confused with the rate of growth of diffusion, $(dm_t/dt.)(1/m_t)$, which can be seen from equation (2.1.2.) to be consistently falling through time. A particularly useful property[2] of β which earns it this title, is that it, <u>alone</u>, defines the time lapse between diffusion reaching any two levels, x_1 and x_2. If $1 > x_2 > x_1$, from (2.1.4),

$$m_t/n = x_1 \text{ when } t_1 = \frac{(\log(x_1/1-x_1) - \alpha)}{\beta}$$

and $$m_t/n = x_2 \text{ when } t_2 = \frac{(\log(x_2/1-x_2) - \alpha)}{\beta}$$

Thus $$t_2 - t_1 = \frac{\log(x_2/1-x_2) - \log(x_1/1-x_1)}{\beta}$$

and, for instance, if $x_2 = .8$ and $x_1 = .2$, then $t_2 - t_1 = \frac{\log 16}{\beta}$.

Whilst the above model is only one of an array of mathematical analyses of epidemics, because of its simplicity, it is the most commonly used in diffusion studies in the social sciences. The analogy to new industrial processes will be discussed presently; the implications of imitative behaviour and bandwagons are fairly obvious.

1. See chapter 6, section 2 of this thesis.

2. This property has encouraged the fitting of logistic curves to real world data. Diffusion speed can thus be measured by a single parameter, even although the growth rate is not constant. This is particularly convenient for comparative studies as will be seen below.

Unfortunately, the stringent assumptions which must be made for such a model to apply are often overlooked and undiscussed. Of particular importance are the following:

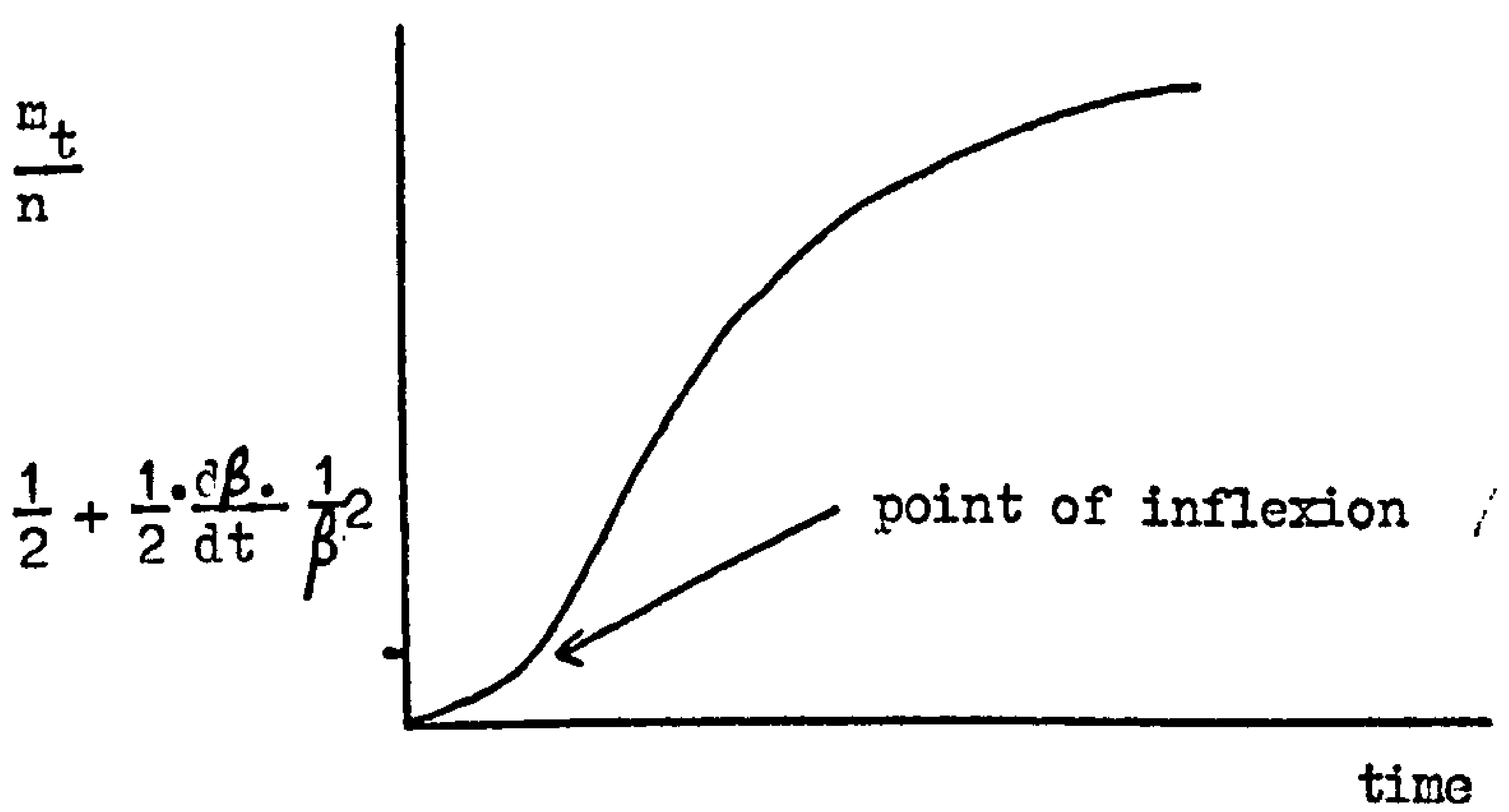
a) <u>once infected, an individual retains the (constant) propensity to infect others.</u> In other words, β is constant. Consider the case, however, of β falling over time, perhaps due to decreasing infectiousness or increasing resistance on the part of the non-infecteds, i.e. $\beta = f(t)$ where $d\beta/dt < 0$ (2.1.5.) Differentiating (2.1.2.) with respect to time and setting d^2m/dt^2 equal to zero, the point of inflexion is given by:

$$1 - 2(m_t/n) + (d\beta/dt)/\beta^2 = 0$$

that is, when $m_t/n = 1/2 + 1/2\,(d\beta/dt)/\beta^2 < 1/2$ (2.1.5a.)

This point of inflexion strongly suggests a positively skewed growth curve.[1]

<u>Figure 2.1.2. A positively skewed S shaped diffusion curve.</u>



b) <u>no individuals withdraw from the population or are cured of the disease.</u>[2]

As an alternative, however, let a constant proportion per period, γ, be cured. Replacing (2.1.2.) with:

$$dm_t/dt = \beta(n - m_t)\,m_t/n - \gamma m_t \quad \text{where } \beta > \gamma \qquad (2.1.6.)$$

Using the method of variation of parameters,[3] the solution of this equation is

1. As an example, let $\beta = a - bt$, then the solution of the revised differential equation is: $m/n = \left\{1 + e^{-(\alpha + at - bt^2/2)}\right\}^{-1}$; this is, indeed, positively skewed for all values of a and b within the relevant range.

2. Clearly, the model would have to be modified to account for bubonic plague!

3. See, for instance, R. Courant, "Differential and Integral Calculus," 2nd edition, Interscience, New York, 1937, vol. 1, pp. 521 - 2.

$$m_t/n = \left\{\beta/(\beta - \gamma) + e^{-(\alpha + (\beta-\gamma)t)}\right\}^{-1} \qquad (2.1.6a.)$$

which is still symmetrical, but in this case the saturation level is reduced to $(\beta - \gamma)/\gamma$.

c) everyone has the same chance of catching the disease. (A necessary, but not sufficient, condition for this and assumption (a) would be a homogeneously mixing population.) The simplest way of relaxing this assumption is to consider two groups, 1 and 2, in the population. At time t, m_{1t} of the n_1 members of the first group and m_{2t} of the n_2 in the second group have adopted. Let group 1 be less susceptible to the disease than group 2, i.e. $\beta_1 < \beta_2$.

Replacing (2.1.2.) with:

$$dm_t/dt = (n_1 - m_{1t})(m_t/n)\beta_1 + (n_2 - m_{2t})(m_t/n)\beta_2 \qquad (2.1.7.)$$

the point of inflexion is given by:

$$1 - 2(m_t/n) - X\,(m_t/n) = 0, \text{ that is, where } m_t/n = 1/(2 + X) \qquad (2.1.7a.)$$

and $$X = \frac{(n_1 - m_{1t})(n_2 - m_{2t})(\beta_1 - \beta_2)^2}{\left[\beta_1(n_1 - m_{1t}) + \beta_2(n_2 - m_{2t})\right]^2}$$

As $X > 0$ for all $\beta_1 \neq \beta_2$, the point of inflexion is reached before 50% diffusion is attained. In this case, a solution to the differential equation seems unattainable. However, as m_t/n will rise from 0 to 1, a point of inflexion at something less than a half, assures a positively skewed S shape.

These examples are sufficient to indicate how fragile the simple logistic solution is - their relevance in the context of industrial innovation will be discussed presently. It should be stressed, finally, that there are many other possibilities which would also invalidate the logistic solution; for instance, if β rises in value over time, a negatively skewed growth curve will result.

2. The inter-industry/innovation approach, (a) the theory.

The above discussion relates primarily to the research considered in this section. The logistic curve is often fitted to diffusion data on different innovations in different industries, and then in a second stage, the $\hat{\beta}$ are used as the dependent variable, in a cross-section analysis, to be explained in terms of the characteristics of the industries and innovations concerned.

By far the most renowned is the contribution of Mansfield.[1] At face value his research does not appear to use the epidemic model outlined above.

Using subscripts i and j for the i th industry and j th innovation, the economic core of his model is contained in the relationship:

$$\lambda_{ijt} = f\left(\frac{m_{ijt}}{n_{ij}},\ \pi_{ij},\ S_{ij},\ \Sigma X_{ij}\right) \qquad (2.2.1.)$$

where λ_t is the proportion of 'hold-outs' (firms not having adopted the innovation at time t that adopt by time t + 1).

$$\text{Thus } \lambda_{ijt} = \frac{m_{ijt+1} - m_{ijt}}{n_{ij} - m_{ijt}} \qquad (2.2.2.)$$

m_t/n is the proportion of firms having adopted by time t; π is the profitability or returns to be gained from adopting the innovation; S is the size of the investment outlay required to install the innovation and ΣX_i a number of other, allegedly less important, and, initially, unspecified variables.

The model is developed by reducing the general function f to a quite specific form. This transformation, however, is conducted on the basis of strictly non-economic manoeuvres: f is approximated by a Taylor's expansion that omits all third and higher order terms, the coefficient of $(m_{ijt}/n_{ij})^2$ in the expansion is set equal to zero because λ_{ijt} is not highly correlated with $(n_{ijt}/n_{ij})^2$ for the innovations in his sample and the period t to t + 1 is assumed to be very small.

$$\text{Thus } f = A_{ij} + \beta_{ij}(m_{ijt}/n_{ij}) \qquad (2.2.3.)$$

$$\text{where } \beta_{ij} = a_1 + a_2\pi_{ij} + a_3S_{ij} \qquad (2.2.4.)$$

(omitting ΣX_{ij} - later to be found non-significant.) A_{ij} represents all other terms in the expansion not containing (m_{ijt}/n_{ij}).

Finally, assuming $\lim_{t\to-\infty} m_{ijt} = 0$, the ensuing differential equation,

$$\frac{dm_{ijt.}}{dt}\ \frac{1}{(n_{ij} - m_{ijt})} = A_{ij} + \beta\frac{m_{ijt}}{n_{ij}} \qquad (2.2.5.)$$

footnote 1. top of next page...

1. E. Mansfield, "Technical change and the rate of imitation," Econometrica 1961, later reprinted in "Industrial research and technological innovation" (1968) op. cit.

has the solution:

$$\frac{m_{ijt}}{n_{ij}} = \frac{1}{1 + e^{-(\alpha + \beta t)}} \qquad (2.2.6.)$$

where α is the constant of integration; the limit condition effectively constraining A = 0.

As can be seen, (2.2.5.) and (2.2.6.) coincide precisely with (2.1.2.) and (2.1.3.) of the epidemic model. Indeed, equation (2.2.1.) and the ensuing mathematics are largely superfluous. One wonders why the model was not stated immediately in the form (2.2.5.) In fact the rationalizations for m_{ijt}/n_{ij}, S_{ij} and Π_{ij} as determinants of λ_{ijt} seem more in keeping with the specific form (2.2.5.) than the general form of (2.2.1.). Mansfield argues that λ_{ijt} will be larger the higher is Π_{ij} because the latter increases "the chance that a firm's estimate of the profitability will be high enough to compensate for whatever risks are involved."[1] λ_{ijt} will be inversely related to S_{ij} as S_{ij} reflects the extent of caution and financing problems associated with potential adoption. λ_{ijt} will be proportional to m_{ijt}/n_{ij} for three reasons: as other firms adopt, (i) competitive pressures mount for non-adopters and bandwagon effects occur, (ii) non-adopters are persuaded to view the profitability more favourably and (iii) the risk attached to adoption declines. Clearly, effects (ii) and (iii) will depend on the values of Π_{ij} and S_{ij}. This might logically lead to a specification such as:

$$\lambda_{ijt} = a_1 (m_{ijt}/n_{ij}) + a_2(m_{ijt}/n_{ij})\Pi_{ij} + a_3(m_{ijt}/n_{ij}) S_{ij} \qquad (2.2.7.)$$

in which each component of the right hand side represents one of the three effects just outlined. (2.2.7.) is, of course, a combination of (2.2.5.) and (2.2.4.)

Thus, to all intents, this model is merely an example of the simple epidemic model presented in the previous section: the choice of independent variables

1. ibid.

is very much in line with the epidemic model and the mathematical assumptions enable Mansfield to attribute an unjustified generality to the initial equation. This does not detract from the ingenious use of the epidemic model in this context. Nevertheless, once the true nature of the model is recognized, certain implicit assumptions emerge which can be seen as crucial to the conclusion that the growth curve will be logistic. These are each analogous to the three critical assumptions (a) - (c) of the epidemic model outlined in the previous section. Specifically, the profitability and cost of the innovation (and thus β_{ij}) must remain constant over time, no firm can become disillusioned and reject the innovation once it has adopted, and all firms must have the same susceptibility to the innovation. A relaxation of any of these assumptions (such as in equations 2.1.5., 2.1.6., 2.1.7.) will rule out the simple logistic form of (2.2.6.). In the following chapters, evidence will be presented that suggests that the first and third of these assumptions, at least, do <u>not</u> hold true for the vast majority of new innovations. Notably, the profitability of the innovation changes over time and, usually, there are economies of scale in adoption, benefiting large firms.

Finally, the theoretical base of the model can be seen as unconvincing. The diffusion process is, after all, the sum total of each individual firm's decision to adopt. An appropriate starting point would surely be a more explicit statement on decision making under uncertainty.

The two other main proponents of this approach can be considered with less discussion. Griliches[1] also uses the logistic to describe the diffusion curve; he postulates, too, that the parameters of the curve will be determined by certain characteristics of the innovations and firms concerned. His choice of the logistic, however, is not based on any underlying model or conviction, "while there are some good reasons why an adjustment process should follow a path which is akin to a logistic, I do not want to argue the relative merits of the various S shapes." Thus, his interest in the

1. Z. Griliches, "Hybrid corn: an exploration in the economics of technological change," Econometrica, October 1957.

curve is only as a tool with which to generate parameters of diffusion to be explained in a second empirical stage.

In subsequent years a number of economists[1] have also <u>assumed</u> logistic curves, justifying their choice either with direct reference to Mansfield's work or with a few passing remarks about epidemics or bandwagons. One exception is Metcalfe.[2] Whilst he still relies on the epidemic analogy, he considers that the assumptions needed to justify the logistic are unlikely to be fulfilled in the case of the diffusion of new industrial processes. Consequently, his choice of growth curve is the logarithmic reciprocal:

$$m_{ijt} = n_{ij}\, e^{-\beta_{ij}/t} \qquad (2.2.8.)$$

This is positively skewed with a point of inflexion at $m_{ijt}/n_{ij} = 1/e^2$. However, this choice is rather arbitrary, (2.2.8.) being employed simply because of its skewness. Whilst good reasons have already been given for a positively skewed growth curve, there is no reason why the degree of skewness should always be the same. But the logarithmic reciprocal implies this and is just as inflexible, in its own way, as is the logistic.

<u>2(b) The empirics.</u>

In spite of the preceeding theoretical criticisms, these models do generate impressive results. Mansfield tests the logistic curve using weighted least squares on the transformation:

$$\log\left(\frac{m}{n-m}\right)_{ijt} = \ell_{ij} + \beta_{ij} t \qquad (2.2.9.)$$

Using data on the diffusion of 12 innovations in the American Iron and Steel, Coal, Rail and Brewing industries, he reports the coefficient of correlation between the dependent variable and time as exceeding .89 in all cases. At face value, then, the logistic does give an adequate description. However,

1. S. Globbermann, "Technological diffusion in the Canadian tool and die industry," York University (unpublished) and P. Swann, "The international diffusion of an innovation," Journal of Industrial Economics, September 1973.

2. J. Metcalfe, "Diffusion of Innovations in the Lancashire textile industry", Manchester School, June 1970.

a number of qualifications may be made. By its very nature, the variable $(m/n)_{ijt}$ must rise monotonically with time, therefore, one would expect a high correlation between any simple transform of it and time. This is especially true given the low number of observations used: an average of only 10 per innovation and in four cases, fewer than 6. Secondly, Mansfield reports significant auto-correlation[1] in at least three cases; quite possibly, this is due to mis-specification of mathematical form. Perhaps a skewed curve might have been more applicable in at least some cases. Thirdly, his data relates only to large firms in the four industries; this is because there were problems in collecting information from smaller firms about the dates at which they adopted the various innovations. This is unfortunate on two counts: (a) only an average of less than 20 firms per innovation are considered, thus constituting only small samples of the total industries, (b) more important, he has effectively reduced the heterogeneity in the four populations. As has been shown already, if large firms are liable to adopt more quickly than small firms,[2] this will tend to produce a skewed curve for diffusion <u>in the population</u>. However, by excluding small firms from the samples, he has restored the homogeneity that is necessary for the logistic to apply. In other words, because the samples are biased, one might doubt the applicability of the logistic to the populations. Agnosticism is perhaps justified on the basis of these three limitations.

In a second stage, Mansfield uses the estimated slope parameters $(\hat{\beta}_{ij})$, as measures of the speed of diffusion in a cross-section analysis based on equation (2.2.4.) In all, six different explanatory variables are used in various combinations, the only two approaching significance being π_{ij} and S_{ij}. π_{ij} is measured as the average pay-out period required, by firms in industry i, to justify typical investments, divided by the average pay-out period actually achieved for innovation j. S_{ij} is the average initial

1. See chapter 6, table 6.4.1, for evidence of auto-correlation found when fitting the logistic to diffusion data for my sample innovations.
2. The next section indicates that this <u>is</u>, indeed, the case.

investment in the innovation j as a percentage of the average total assets of the firms in industry i.

The best explanation achieved is:

$$\hat{\beta}_{ij} = \begin{bmatrix} -0.29 \\ -0.57 \\ -0.52 \\ -0.59 \end{bmatrix} + \underset{(.015)}{.530}\,\Pi_{ij} - \underset{(.014)}{.027}\,S_{ij} \qquad R^2 = .99 \qquad (2.2.10)$$

where figures in round brackets denote estimated standard errors.

The four alternative intercept terms correspond to the four industries. Although Mansfield is not explicit, these have been estimated, presumably, using dummy variables.[1]

Mansfield draws four main conclusions:

(i) "(the equation) represents the data surprisingly well."

(ii) the coefficients of Π_{ij} and S_{ij} have the expected signs.

(iii) both coefficients are significantly different from zero at the 5% level. (apparently, S_{ij}'s coefficient is significant only if an extreme observation is included.[2])

(iv) the differences in the size of the intercept terms are not inconsistent with more concentrated industries being slower at diffusion.[3]

These results must be judged in light of the small sample size (12 observations) and the fact that six explanatory variables are used (including the four dummies) but nevertheless, the overall fit is remarkable, given the nature of the dependent variable.

It would be helpful to know how much of this excellent explanation is due to the dummy intercept terms. One can conceive of β varying due to differing characteristics of (a) the innovations and (b) the industries concerned. From this equation, we know that at least one innovation characteristic (Π_{ij}) is crucial, but we are ignorant of which industry-level characteristics might be important. The differential intercepts establish that there <u>are</u>

1. At any event, each of these intercepts has been estimated on the basis of only 3 observations; there being 3 innovations for each industry.
2. In fact, even including this observation (as in the quoted equation), the coefficient is surely insignificant.
3. He ranks the industries according to how competitive he thinks they are, and finds a rank correlation coefficient between this ordering and that of the intercept terms of .80. As he, himself, concedes, this is hardly a very strong test, however.

inter-industry differences after normalising for innovation characteristics. Whether these differences are significant, however, is not known. Further, there is no possibility of discovering their causes, given the small industry sample size.

Unfortunately, Griliches' results can shed no light on this matter. His data refer to the diffusion of an agricultural innovation - hybrid corn - in 31 different American states between 1932 and 1956. He measures diffusion as 'the percentage of all corn acreage planted to hybrid seed' i.e. overall, as opposed to inter-firm, diffusion. The logistic is fitted to the data in much the same way, except that Griliches does not impose saturation levels (n_{ij}) but uses the results to decide between a number of alternatives[1]. Again, the logistic fits the data well - in none of the 25 cases does R^2 fall below .89 - but as Durbin-Watson statistics are not reported, we can not tell whether alternative time curves might have performed better. In his second stage, Griliches attempts to explain the inter-state variance in the three estimated parameters in a cross-section analysis similar to equation (2.2.10). Whilst the fit achieved in this stage is not as high as in Mansfield's second stage (e.g. an R^2 of .5 - .6 is typical in explaining $\hat{\beta}_{ij}$), a number of significant determinants are detected. Almost without exception however, these variables relate to the relative profitability of using hybrid seed and are usually related to inter-state geographical differences such as fertility of soil. These results have led to interpretations such as 'Griliches's study . . . . shows that the behaviour of both farmers and hybrid-seed producers were firmly grounded in expectation of profit,' [2] and confirms Mansfield's finding of the importance of π_{ij}. On the other hand the role of inter-industry variables such as industrial structure, growth of market etc. remains unexamined.

1. That is, in fitting the equation (2.2.9), he experiments with different values of n and selects for each state the n which maximises the R^2. This means of course, that he has three parameters for each logistic as opposed to only two for Mansfield, and correspondingly, one degree of freedom less.

2. N. Rosenberg, 'The economics of technological change.' (Penguin Readings, London 1971) p. 209.

This omission is also unavoidable in Metcalfe's study of the diffusion of three innovations in the Lancashire textile industry. Of course, by studying only one industry, he effectively removes inter-industry differences and concentrates on innovation characteristics. Using the alternative S shaped curve(equation 2.2.8), he fits the transformation:

$$\log n_{ijt} = \log n_{ij} - \beta_{ij}\left(1/t\right) \qquad (2.2.11)$$

and finds a satisfactory fit in each case, although, once again, no Durbin -Watson statistics are quoted. Obviously, no systematic cross-section explanation of $\hat{\beta}_{ij}$ is possible given only three observations; but it is noticeable that the innovation recording the slowest speed of diffusion is also the most expensive and least profitable.

In summary then, none of these three studies establish that the chosen S shape necessarily yields the best explanation of the time series data; but, on the other hand, all three present evidence to suggest that the characteristics of the innovations do affect their speed of diffusion. The role of industry-level variables remains largely unexplored.

A note of dissonance.

Even though Mansfield's and Griliches's studies were published more than a decade ago, there has been very little critical analysis of their work. One notable exception is provided by Gold, Pierce and Rosseger[1] in their own study of the diffusion of 13 major process innovations in the U.S. They claim that there is such a diversity of variables which affect the diffusion of individual innovations that it is almost pointless to build a general model of diffusion as did Mansfield. For instance, by product coking diffused only very slowly initially as there were plentiful supplies of coking coal and organic chemicals; on the other hand, beneficating and pelletizing spread quickly from the first introduction due to shortages of high grade iron ore; machine cutting of coal, like by product coking, did not 'catch on' very quickly, but this time because there were a large number of variants of the technique available on the market which led to confusion

1. B.Gold, W.Pierce, G.Rosseger, "Diffusion of major technological innovations in U.S. Iron and Steel," Journal of Industrial Economics, July 1970.

and uncertainty amongst potential consumers. Indeed, they cite 'special circumstances' for all of the innovations they study. This seems to be an unnecessarily pessimistic view. In any cross-section study, in virtually any area of economics, there are special or random influences; the point of model building is to investigate the importance of general factors; it may be, of course, that special factors dominate, but that, surely, is a matter of empirics. As it happens, an examination of the 'special factors' in this case suggests that a number of them reduce to differences in the profitability of the innovations, on the one hand; or different rates of change of profitability on the other. (Only the latter are beyond analysis in Mansfield's model, but there is no reason why they could not be incorporated into a general model.)

A second broad criticism made of Mansfield's approach is his use of ex-post profitability as a determinant of the speed of diffusion. The arguments are that: many business decisions are the results of animal spirits, expected profitability is a more meaningful concept anyway in this case and that the profitability of any innovation will not be constant over the diffusion period. Whilst these are all valid criticisms (hopefully to be partly answered by the model developed in this thesis,) Mansfield's answer, one suspects, would be pragmatic - namely, that his specification was dictated by availability of data.

On a more positive note, Gold et al do have interesting observations regarding the role of the demand facing the adopting industries. Of the 5 innovations introduced in slow growth periods (i.e. where demand for the industry's product over the first 15 years of the innovation's life was slow or static), none diffused rapidly. Of the 8 innovations introduced during periods of rapid growth, 3 diffused rapidly. The (rather heroic) conclusion drawn is that industry growth may be necessary but not sufficient for fast diffusion.

3. Inter-firm differences (a) theory.

The central aim of this body of research is the explanation of differences between firms in the speed with which they adopt the same innovation. Again, the innovator is Mansfield who postulates[1] the following relationship:

$$d_{ij} = Q_i . H_{ij}^{a_{i2}} S_{ij}^{a_{i3}} G_{ij}^{a_{i4}} \Pi_{ij}^{a_{i5}} A_{ij}^{a_{i6}} L_{ij}^{a_{i7}} T_{ij}^{a_{i8}} e^{\epsilon_{ij}} \quad (2.3.1)$$

where d_{ij} is the number of years the jth firm waits before beginning to use the ith innovation, S_{ij} the firm's size, H_{ij} a measure of the profitability of its investment in the innovation, G_{ij} the firm's rate of growth, Π_{ij} the firm's profitability, A_{ij} the age of its president, L_{ij} a measure of its liquidity, T_{ij} the firm's profit trend. It should be stressed that this part of his work is quite independent of his epidemic model discussed in the previous section. Nevertheless, it can be seen that the diffusion growth curve for innovation i is merely an aggregation of d_{ij} over all firms in the relevant industry, therefore, one might expect a common theoretical thread in the two fields. In fact, not only is the common thread missing, but also (2.3.1) is based not on any conscious model of firm behaviour but rather, on ad-hoc theorising. There are, however, broad hints of a behaviouralist view: declining profits are seen as activating search and older presidents are believed to be more conservative (i.e. $\partial d/\partial T < 0$; $\partial d/\partial A > 0$). Similarly, the role of the technological characteristics are hinted at but not developed. For instance, poor liquid asset ratios and low profits will hinder adoption of multi-million pound innovations but probably not of relatively cheap innovations. One might expect, therefore, a_{i5} and a_{i7} to depend on the type of innovation, but this point is not discussed.

The two most important variables (both empirically and theoretically) are S_{ij} and H_{ij}. d_{ij} is likely to be inversely related to S_{ij} on three counts. First, the costs and risks of early adoption are more easily borne by large firms. Second, because of their size, large firms are more likely to have, at any point in time, a greater probability of needing to replace old equipment, and thus to the extent that the innovation is embodied in new equipment, large firms have greater opportunities to adopt early, on average. Third, again purely because of their size, larger firms

1. E. Mansfield, "The speed of response of firms to new techniques," Quarterly Journal of Economics, May 1963.

are likely to encompass a wider range of operating conditions than smaller firms. As some innovations, initially, have only limited applicability, there is more likelihood that large firms will have the appropriate operating conditions needed for adoption of the new innovation in its early years. For instance, certain innovations in paper making were initially only applicable to the production of certain special types of paper. Other things being equal, large firms were more likely to include these papers in their range of products, as opposed to smaller firms specialising in only one or two paper types.

The profitability of the innovation for firm j, H_{ij} presents measurement problems as it is likely to be determined by a number of characteristics of the firm, such as its product mix, quality of inputs etc. A statistical possibility, not acknowledged by Mansfield, is that H_{ij} may be collinear with S_{ij} (and perhaps G_{ij} and L_{ij}) if there are significant returns to scale in the adoption of the innovation.

Finally, the multiplicative form of (2.3.1.) is rationalised on the grounds of interdependence of influence of the independent variables and the desired possibility of negative second order derivatives.

Most other research in this area follows Mansfield's ad-hoc theorising and empirical methodology quite closely. A number of authors, however, have placed more emphasis on 'attitudional' or 'informational' variables, whilst still retaining firms' economic characteristics in their analysis. These variables will be discussed individually below. As a group, however, their inspiration is to be found in the sociologist Rogers'[1] survey of diffusion studies in the other social sciences. Apparently, the individuals who appear to be quicker to adopt new ideas, techniques etc., tend to be young, affluent, opinion-leaders, non-traditional and cosmopolite (i.e. mix with individuals outside of their own groups or even countries.) Such an approach does present problems. Often, such characteristics are difficult to measure without resorting to the construction of arbitrary indexes (perhaps

1. E. Rodgers, (1962), op. cit.

based on questionnaire answers.) Sometimes circular results arise e.g. if non-traditional, opinion leaders are found to adopt earlier, the only conclusion which can be drawn is that progressive firms adopt new innovations quickly. Further, even if firm decision making is dominated by one man or a small group whose personal characteristics play a significant role, such variables, given the above data problems, only merit inclusion in the analysis if their effects are independent of the economic characteristics of the firm, such as its size, growth, profits etc.

3(b) Empirics.

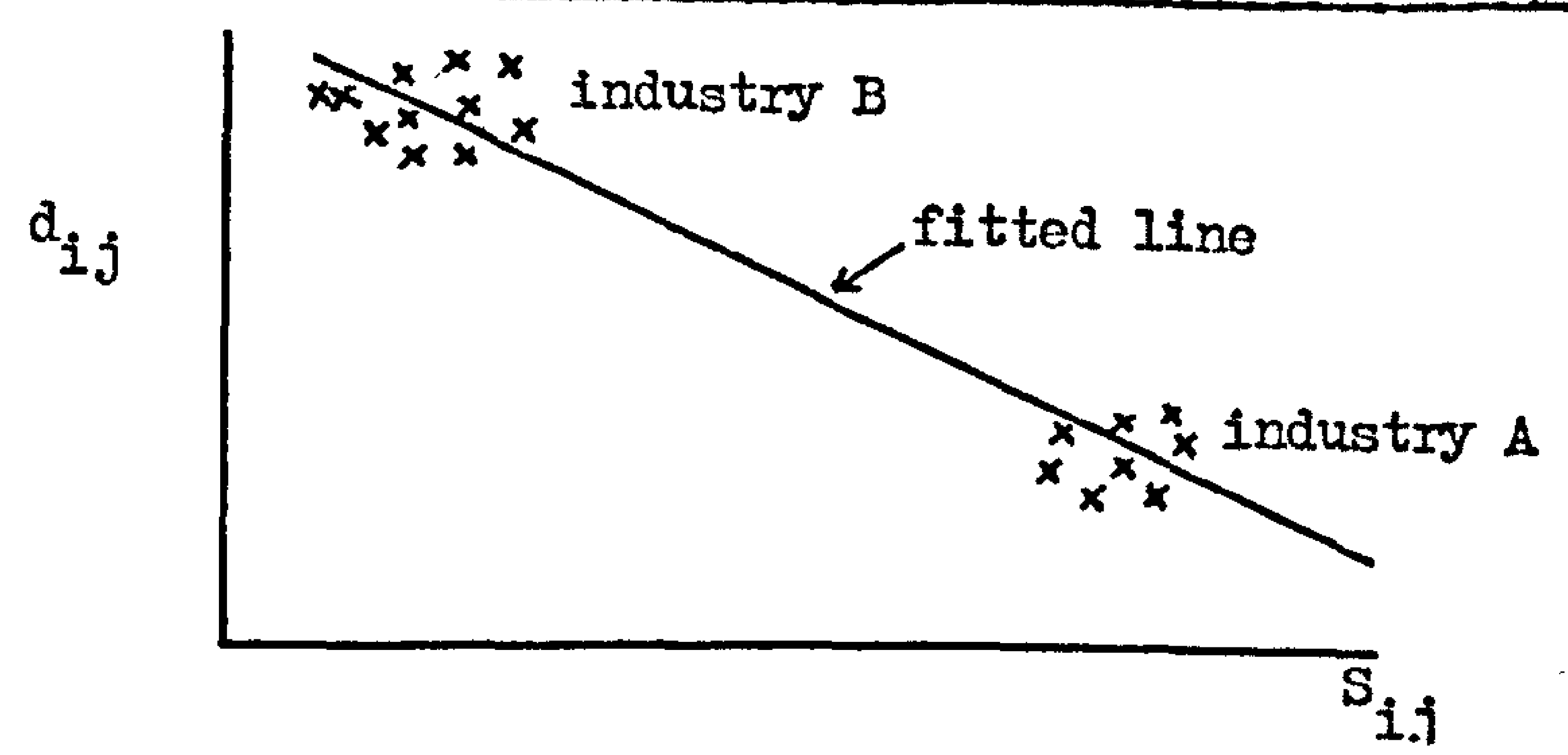
Empirical findings in this area are unimpressive and even the few significant results must be viewed in the light of serious statistical qualifications. One suspects that the potentially large number of technological factors affecting the date of any one firm's adoption places strict limits on the degree of empirical success in this particular area.

Mansfield applies his model to data for 167 firms adopting 14 different innovations in various American industries. Only S_{ij} and H_{ij} attain statistical significance and the latter for only 2 of the 14 innovations. Unfortunately, H_{ij} cannot be measured directly and so proxy measures are used. For instance, the profitability in adopting a continuous mining machine varies directly with the proportion of the firm's output derived from high seams, consequently this proportion for each firm is used as proxy for H. Unfortunately, profitability probably depends upon many other unknown or unmeasurable factors which obviously restrict the explanatory power of his model. At any event, even such crude measures are available for only 5 innovations.

S_{ij} is consistently significant, however, regardless of which other variables are included in the equations; $\hat{a}_3$ varies between -.03 and -1.53 with Mansfield's preferred equation yielding $\hat{a}_3 = -.4$.

His results are rather ambiguously presented (no R^2 are mentioned) but quite obviously the model is tested against grouped data for all industries and innovations. However, whilst Q_i and a_{i2} are free to vary between innovations, a_{i3} is constrained to be equal for all i. This might be questioned on theoretical grounds: the importance of firm size would surely vary depending on the type of innovation and/or industry. Statistically, the assumption introduces ambiguity in the interpretation of $\hat{a}_3$: one cannot be sure that $\hat{a}_3$ reflects inter-industry, as opposed to inter-firm size, effects. As an extreme example, consider 2 industries adopting different innovations: industry A comprising a few large firms adopts its innovation quicker than industry B, which comprises many small firms. However, within each industry, size of firm has no effect on adoption date. The hypothetical scattergram might be as in the figure (2.3.1.) The O.L.S. line fitted to such data is, of course, downward sloping even although, within each cluster (industry), there is no discernible relationship,

Figure 2.3.1. The confusion of industry and firm size.



As Mansfield studies 14 innovations in 4 different industries, such ambiguity could easily have been avoided by letting a_3 vary for different innovations and industries. Inter-industry size effects could be tested using t tests on the $\hat{a}_{i3}$. As his results stand, however, one must remain agnostic on the role of firm size, and for that matter, industry size.

The results of other empirical research are equally disappointing, being characterised by low R^2, insignificant variables and rather suspect specification of the dependent variable.

Nabseth,[1] using data for Swedish firms adopting 6 different innovations, employs up to nine explanatory variables, of which four approximate roughly to Mansfield's S_{ij}, H_{ij}, G_{ij} and I_{ij}, the other five being 'attitudional'

1. see *1. at bottom of following page (20)

variables. Of the first group, only S_{ij} and H_{ij} are ever significant and then only for 2 and 1 of the 6 innovations respectively. Of the second group, two (INF_{ij} and B_{ij}) are significant for 4 and 3 innovations. INF_{ij} measures the date at which firm j claims to have first heard of innovation i, and B_{ij} is an arbitrary index for which firms are given ratings between 0 and 16, depending on how quickly they adopted past innovations. In other words, there is some evidence to suggest that firms adopt more quickly, the sooner they know of the existence of the innovation and the more progressive they are.

Hakonson[1] uses the same set of variables in his study of the diffusion of Special Presses in three countries. As is true for Nasbeth's regressions, R^2 in the region of .4 to .6 are attained. S_i and H_i are both significant with expected signs in 2 of the 3 countries.

Smith's [2] results (again studying a single innovation in a number of different countries) are even more disappointing. Using much the same array of explanatory variables, R^2s of .1 are typical and only an arbitrary index reflecting the extent of firms' vertical integration attains significance.

Each of these three authors face a problem not encountered by Mansfield – they are studying innovations which have not yet diffused 100%; consequently, they do not have observations on their dependent variable d_{ij} for all firms. Rather than discard non-adopting firms from their samples, they allocate to such firms an arbitrary adoption date in the future.[3] The unfortunate result of this step is almost certain to be biassed estimates. This can be seen using a simple example.

Suppose that the true relationship is $d_j = \alpha + \beta X_j + \epsilon_j$ (2.3.2.) which would yield a scattergram as in figure (2.3.2.) (a). However, if

*1.(footnote carried over from bottom of page 19.) L. Nabseth, "The diffusion of innovations in Swedish industry," in "Science and technology in economic growth," ed. B.R. Williams, McMillan, London, 1973.

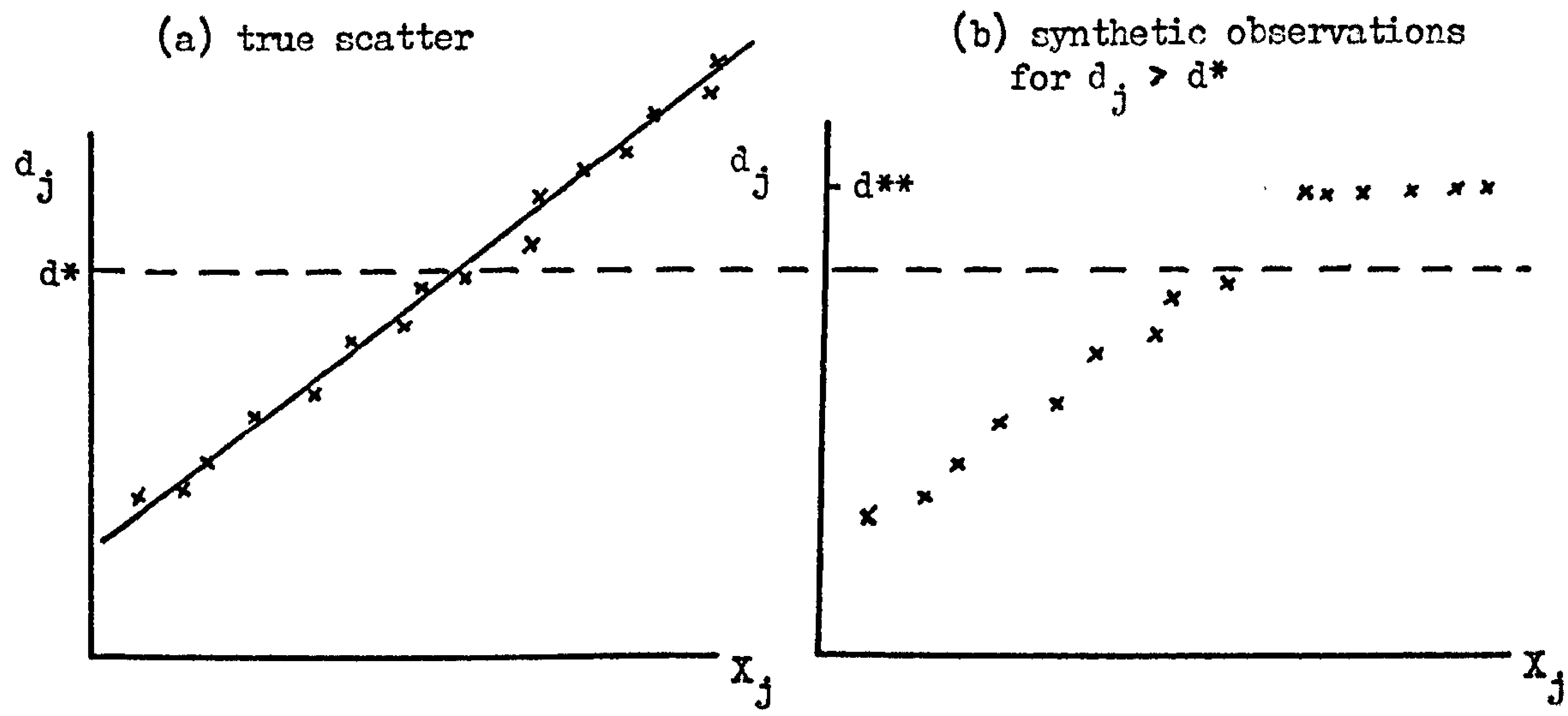
1.S.Hakonson,"Special presses in papermaking," Ch. 4 in "The diffusion of new industrial processes; An international study," ed.L. Nabseth and G.Ray, Cambridge University Press, 1974.

2. R.Smith, "Shuttleless looms," chapter 10, ibid.

3. Nabseth and Hakonson assumed that non-adopters would adopt in 1975, whilst Smith assumed adoption in 1980.

viewed at time d* (i.e. the time of the research) no observations are

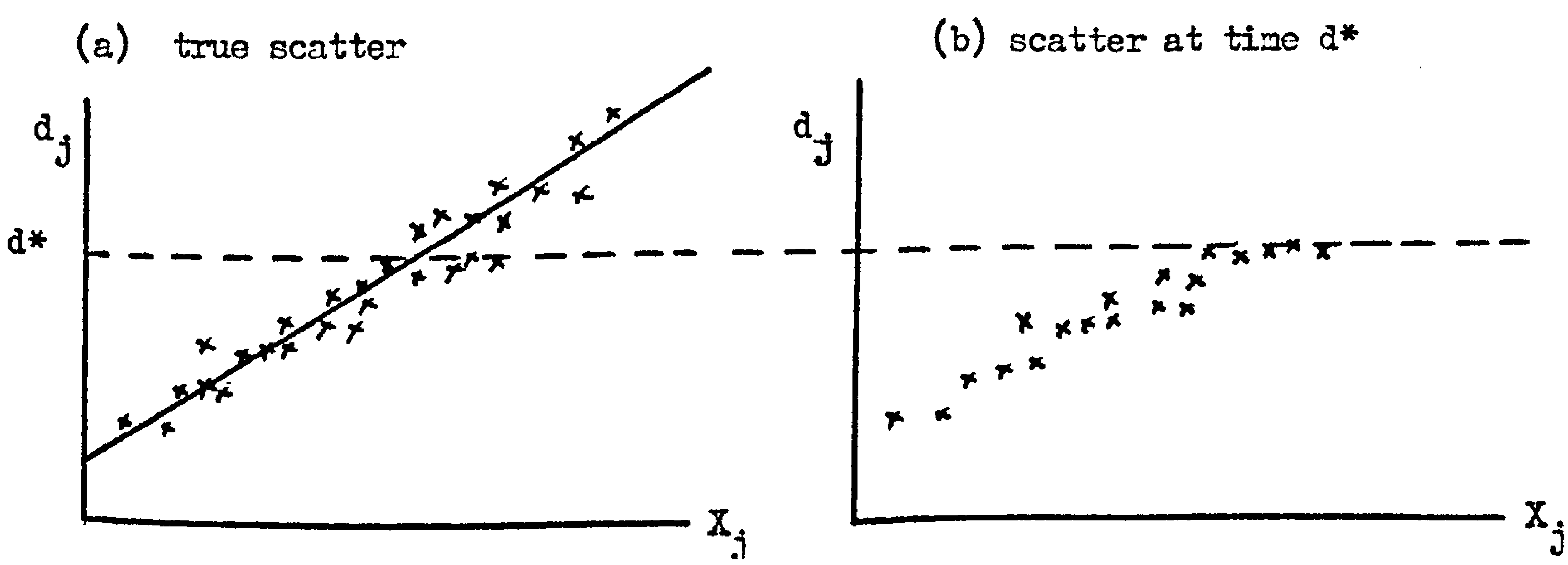
figure 2.3.2. The cause of biassed estimates of β.



available for non-adopters (having d > d*), these firms are thus attributed with d_j = d**. Clearly, if a line were to be fitted to the scatter in 2.3.2, $\hat{\alpha}$ and $\hat{\beta}$ would be biassed estimates of α and β. As the figure has been drawn, $\hat{\beta}$ would be biassed downwards, but this need not always be the case.

Whilst this problem invalidates the results quoted above (excluding Mansfield's), a 'correct' way of incorporating non-adopting firms into the analysis is not immediately obvious. It would be equally suspect to omit the non-adopters from the analysis altogether: the sample would then be unrepresentative. The possible consequences can be seen in figure (2.3.3.) Again, (2.3.3a.) portrays the true scatter and line, (2.3.3(b)) portrays the position omitting non-adopters at time d*.

figure 2.3.3: An alternative cause of bias.



Again, biassed estimates of α and β would ensue. Intuitively, the bias would appear to be less important the smaller is var (ϵ_j).

This dilemma is of direct relevance to the empirical methodology to be pursued in this thesis. For nearly all the innovations in my sample, diffusion is not complete at the time of writing. These problems appear to rule out, therefore, <u>this sort</u> of empirical work on inter-firm differences.[1]

Finally, even the tentative firm-size influence implied by the above research is faced with the conflicting, non-statistical evidence of Adams and Dirlam.[2] In their study of the diffusion of the Basic Oxygen Process in the American steel industry, they note that the first innovator and early adopters were all small firms and that all of the major firms in the industry held back from adopting for a number of years. Their rationale of this rests on the hypothesis of oligopoly fostering implicit collusion: 'it may well be that the structural and behavioural characteristics of oligopolised industries <u>prevent</u> the dominant firms from pioneering.' Small firms in such an industry can adopt a new innovation and even pass on the cost savings in lower prices without necessarily destroying the status quo. On this basis it would be imprudent to suppose that firm size need always act in the same way in this respect, regardless of industry structure.

4. International Comparisons.

The empirical methodology in this area has been somewhat less uniform than in either the inter-industry or inter-firm approaches. The very nature of the comparison permits a different emphasis than is necessary in, say, Mansfield's work; as the <u>same</u> innovation is studied in different countries, the differential diffusion rates are due, more probably, to the characteristics of the industries and countries concerned. Indeed, certain institutional international differences do emerge which one suspects are a major cause of

1. In fact a way round the problem might be to fit the equations subject to inequality constraints for non-adopters. However, as the estimated Quasi-Engel curves of chapter 7 provide an alternative but equivalent approach, this possibility will not be pursued.

2. W.Adams and J.Dirlam, "Big steel invention and innovation", Quarterly Journal of Economics, May 1966.

the diversity in approaches.

Nevertheless, the Mansfield methodology is followed in a number of instances. Swann[1], for instance, fits logistic curves to data on the diffusion of synthetic rubber in 12 countries during the post-war period. In a second stage, he uses the parameters of the fitted logistics as the explained variables in a cross-section analysis using country-level explanatory variables such as the growth in output, the level of rubber imports, the level of rubber exports and the production of rubber per capita. Both his curve fitting and cross-section analysis record high R^2 and significant variables.[2] His rationalisation of the logistic and explanatory variables is brief and along familiar lines;[3] again there is no fully worked economic model.

The major analysis of inter-country differences has been produced by an international consortium of research institutes. The diffusion of ten different innovations in up to six different countries is studied; each institute being responsible for the analysis of one innovation[4]. (which in turn are allocated separate chapters in the report on the project.) An inevitable result of this demarcation is that different methodologies have been used in each case and general conclusions are difficult. Having said this, a concensus view does tend to emerge. Whilst a number of the authors concede the likelihood that the time path of diffusion is S shaped (usually rationalised with reference to the epidemic model), curve fitting is not attempted in most cases. This is surely quite acceptable <u>if one believes that there is no economic, non-trivial rationale for the S shape</u>.

1. P.Swann, 'The international diffusion of an innovation,' Journal of industrial Economics, September 1973.

2. However, his successful explanation of inter-country cross-section differences in the speed parameter β, depends on the use of an explanatory variable measuring the date at which the innovation 'took off' in the country concerned, defined as the date at which the curve predicted 10% diffusion i.e. $(-2.2-\hat{\alpha}_i/\hat{\beta}_i)$.In other words, $\hat{\beta}_i$ is regressed against a variable which is directly proportional to $\hat{\beta}_i$. Not surprisingly a highly significant positive coefficient emerges.

3. See the rationale in the previous section for similar variables at the firm level.

4. Except N.I.E.S.R. who were responsible for five.L.Nabseth and G.Ray, op.cit., 1974.

If the S-shape is due to a physical quasi-sociological law, then curve fitting is a purely descriptive procedure, designed only to generate some empirically useful measure of the speed of diffusion (the slope parameters, β, in the case of the logistic.) In line with virtually all empirical work on diffusion, the underlying theory of the firm generally appears more behavioural than neo-classical, with a number of authors emphasising the role of management factors (various indices of attitudes and information receptiveness are constructed.) Generally, however, inter-country differences are explained in terms of three sets of variables. Most popular are measurements or proxies for the <u>profitability of the innovation</u> in different countries. Oppenlander attributes differential diffusion rates with respect to Numerically controlled machine tools to differences in labour costs; Meyer and Herregat emphasise the role of various factor prices (e.g. scrap metal, labour, capital) in the diffusion of Basic Oxygen Steel-making; Gebhardt attributes prime importance to the compatibility of the country's upstream steel-making processes to the continuous casting technique and so on. Second, and closely related, <u>technological and institutional differences</u> are mentioned in a number of cases. Davies, Smith and Lacci introduce the concept of a Technological Ceiling - different countries are limited to differing extents by the nature of their clay in the extent to which Tunnel Kilns may be adopted in brick-making; Gebhardt uses a similar concept in his analysis of Continuous Casting. Ray highlights the importance of legal restrictions to the diffusion of Gibberellic Acid in Brewing and of licensing agreements to the diffusion of the Float Glass technique in glassmaking. Third, some authors attempt to explain differentials in terms of more <u>conventional economic industry characteristics</u> : growth of market, overall size of the industry and its firms, typical age of existing equipment etc. (Surprisingly little attention is focussed on the intensity of competition.)

Perhaps the major success of the project as a whole is in establishing the magnitude of the empirical and theoretical task facing research in diffusion.[1] The typical innovation is <u>not</u> necessarily suitable for all

1.To this extent, partial support is offered for the conclusions of Gold et al as described earlier firms in the relevant industry, its profitability depends on a host of special factors and may vary widely across firms in the same industry; it is sometimes difficult to specify what the innovation is - its specifications may vary across countries and over time.

5. Stock Adjustment Models.

Nearly all of the work surveyed so far can be seen as a broad continuation of various aspects of Mansfield's seminal contributions of the early sixties. However an alternative methodological approach has also developed with rather less acknowledgement in most surveys of diffusion: the stock adjustment model. Particularly interesting are the studies of the spread of computer usage in the U.S. (Chow[1]) and the U.K. (Stoneman[2]).

Chow and Stoneman postulate that the growth in computer usage in any time period is proportional to the extent to which the actual stock at the beginning of the period falls short of the equilibrium stock. Both employ two alternative forms to test this hypothesis:

$$\frac{dm_t}{dt}\frac{1}{m_t} = \beta(n_t - m_t) \quad \text{or} \quad \frac{dm_t}{dt} = \beta(n_t - m_t)\; m_t \qquad (2.5.1.)$$

$$\text{and} \quad \frac{dm_t}{dt}\frac{1}{m_t} = \beta(\log n_t - \log m_t) \quad \text{or} \quad \frac{dm_t}{dt} = \beta(\log n_t - \log m_t)m_t \qquad (2.5.2.)$$

(where continuous time is used here to indicate the similarities with the epidemic model.)

In other words, the increase in the stock of computers will be determined by the level of the stock (rationalised on familiar grounds of competitive pressures and as a proxy for the quality of information about computers), and the shortfall existing between actual and equilibrium stock.[3] The only difference between (2.5.1.) and (2.5.2.) lies in the precise formulation

1. G.Chow, "Technological change and the demand for computers," American Economic Review, 1967.
2. P.Stoneman, "On the change in techniques - a study of the spread of computer usage in the U.K., 1954 - 70," Ph.D. Thesis, Cambridge (1974).
3. As such, their model may be seen as an extension of the work on time trends in input-output coefficients. See, for instance, K.Wigley, "The demand for fuel, 1948-75", Vol. 8 of "A programme for Growth", Cambridge, 1968, pp. 10 - 14.

of the stock adjustment mechanism. The first form is, of course, the differential equation of the standard logistic; the second is the differential equation of the Gompertz curve which is a skewed S-shape growth curve having a point of inflexion at $m_t/n_t = .37$ (as opposed to .5 for the logistic.) Both authors prefer the Gompertz on the basis of its superior subsequent empirical performance.

Apart from the fact that the estimating forms used are the differential equations, rather than their solutions, neither formulation differs radically from the epidemic model at first sight. However, n_t is also hypothesized to increase with time. In both cases, $n_t = f(X_t, p_t)$ where X_t is G.N.P. and p_t the relative price of computers[1]; $\frac{\partial n_t}{\partial X_t} > 0$, $\frac{\partial n_t}{\partial p_t} < 0$. In other words, the equilibrium stock changes over time as G.N.P. increases lead to increased scope for computers, and as relative price decreases lead to increasing cost savings effected by computer usage. Substituting in a specific expression for n_t, the estimating Gompertz equation is given as:

$$\log m_t - \log m_{t-1} = \beta(b_0 + b_1 \log p_t + b_2 \log X_t) - \beta \log m_{t-1} \quad (2.5.3.)$$

This is estimated for time series data on the post-war spread of computers. Chow estimates β at about .25, but finds only $\log m_{t-1}$ to be significant. Stoneman's results are also a little disappointing at this stage. But as an extra refinement, Stoneman allows β, the coefficient of adjustment, to vary with certain economic variables. Briefly, his argument is that the extent to which the actual stock is adjusted towards the equilibrium stock will vary, depending on the extent to which firms are motivated to search for new methods of achieving their goals and also the efficacy of this search. This behaviouralist stance suggests determinants of β such as the growth of profits, costs and output (determining the goal achievement pressures) and the level of profits, the price of computers and the actual stock (determining the outcome of their feasibility and evaluation processes.)

1. Both authors go to considerable lengths to allow for quality changes in computers in measuring m_t and p_t.

Unfortunately, this refinement involves estimating equations that are under-identified - mainly because n_t and β_t are both now variable. Attempts are made to circumvent the problem by using ancilliary information to estimate n_t, for instance. Although growth of sales appears to have some explanatory power, generally, results remain inconclusive.

Abstracting from the generally unexceptional results of both Chow and Stoneman, their work represents an important deviation from the norm in diffusion studies. First, the emphasis is switched : the estimated growth curve is not merely a <u>means</u> of generating observations to be used in cross-industry comparisons but is, instead, an empirical <u>end</u> in itself. This is only worthwhile, of course, because they, particularly Stoneman, invest the differential equation with strong economic meaning based on an explicit theory of decision making. On the other hand, this does not rule out the possibility of using such an approach for cross-industry comparisons, although there is, no longer, a unique measure of the rate of diffusion. Second, the skewed Gompertz curve describes their time-series data better than the symmetrical logistic. Third, they measure diffusion by the number of computers rather than the number of firms having adopted a computer - this is dictated by the nature of their data as well as by the theory. There is, unfortunately, a theoretical drawback; their model does not differentiate between the initial decision by any firm to adopt a computer for the first time and the later decisions to add to its existing stock of computers.[1] Yet one can envisage different sets of variables influencing these two quite different decisions. To give an obvious example, uncertainty will be important in the initial adoption decision, but less so when adopting further computers; on the other hand, the age distribution of a firm's existing capital stock is likely to be much more decisive to the second decision than to the initial decision.

Nevertheless, this approach could be used to study inter-firm diffusion. However, such a model will not be pursued in the following study, partly because the Gompertz is just as restrictive, in its own way, as is the

1. In Mansfield's terminology, inter-firm and intra-firm diffusion are not differentiated.

logistic and partly because β, the adjustment coefficient, is still specified in a rather ad-hoc fashion.

6. The Diffusion of Consumer Durables.

Surprisingly the study of the diffusion of new products (mainly consumer durables) has developed almost independently of the various studies already mentioned on new processes. Whilst the economic agents at the centre of the analysis are different (individual consumers as opposed to firms), there are sufficient common factors, such as the role of uncertainty and information flows, to suggest that both areas would benefit from cross-polination.

No attempt will be made here to survey this area comprehensively,[1] but one particular technique which has been used will be elaborated, given the central role it will play in chapter 5. Probit analysis has long been used in Biology and other sciences to analyse such things as the efficacies of different dosages of poisons in exterminating insect populations. The technique has been employed with some success in the study of the diffusion of various consumer durables, notably by Cramer, Aitchison and Brown, and Bonus.[2] Only an exposition of the basic framework is submitted here, whilst it can not be attributed to any one of these authors alone, it does represent the common spirit of their work. No reference will be made to their empirical work.

The central assumption is that an individual consumer will be found to own the new product at time t if his income, y_{it}, exceeds some critical level y^*_{it}. This critical or tolerance income may be thought of as representing

1. For a fairly comprehensive survey, see A. Bain, "The growth of T.V. ownership in the U.K. since the war," Cambridge University Press, 1964, chapter 2.

2. J.S.Cramer, "Empirical econometrics," North Holland, Amsterdam, 1969, chapter 3; J. Aitchison and J. Brown, "The lognormal distribution," Cambridge University Press, 1957, Chapters 7 and 12; H. Bonus, "Quasi-Engel curves, diffusion and the ownership of major consumer durables," Journal of political Economy, May / June 1973. See also, F.G.Pyatt, "Priority patterns and the demand for household durable goods", Cambridge University Press, 1964, for an extension of Probit analysis which generates, as a special case, a stock adjustment model not unlike those discussed in the previous section.

the tastes of the consumer which, in turn, may be related to any number of personal or economic characteristics (generally excluding income however). At any event, y^*_{it} is usually regarded as the product of a large number of random influences. As such, the multiplicative form of the Central Limit theorem suggests that y^*_{it} may be lognormally distributed across consumers. The third assumption usually made is that income, itself, is also lognormally distributed across consumers.

These three assumptions may be written as follows:

$$P\left\{q_{it} = 1 \;/\; y_{it}\right\} = P\left\{y^*_{it} \leq y_{it}\right\} \qquad (2.6.1.)$$

that is, the probability that consumer i owns the innovation at time t, given an income of y_{it}, is equal to the probability that his actual income is not less than his tolerance income.

$$y_{it} \text{ is } \Lambda(\mu_t, \sigma^2_t) \qquad (2.6.2.)$$

$$y^*_{it} \text{ is } \Lambda(\mu^*_t, \sigma^{*2}_t) \qquad (2.6.3.)$$

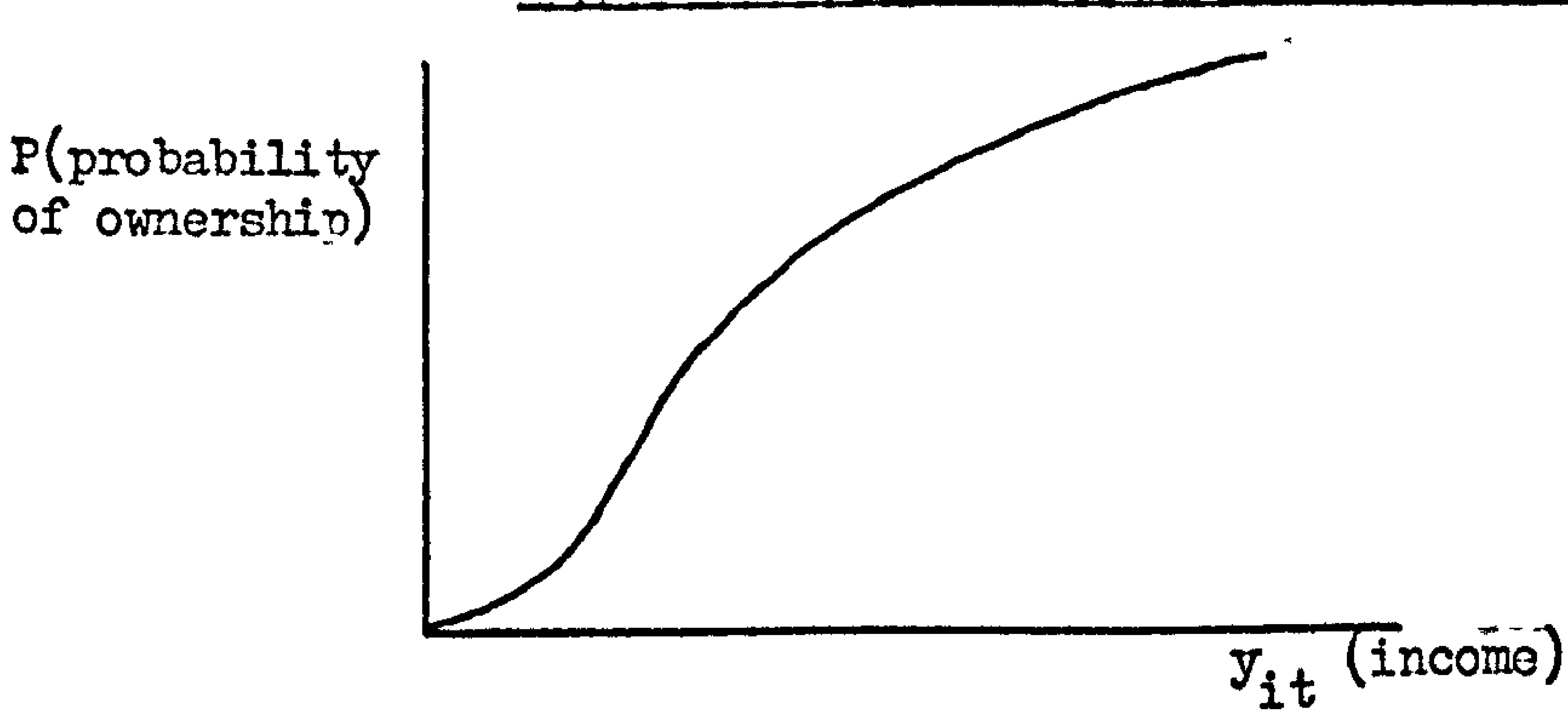
that is, $\log y_{it}$ is normally distributed with mean μ_t and variance σ^2_t and $\log y^*_{it}$ is also normally distributed with mean μ^*_t and variance σ^{*2}_t.

This standard model generates predictions as to the shape of the so-called Quasi or Pseudo Engel curve which relates the probability of ownership to income:

$$P\left\{y^*_{it} \leq y_{it}\right\} = \Lambda(y_{it} \mid \mu^*_t, \sigma^{*2}_t) \qquad (2.6.4.)$$

where the second expression represents the proportion of the tastes distribution having $y^*_{it} \leq y_{it}$. Therefore P rises monotonically with y_{it}. In fact, the Quasi-Engel curve is simply a positively skewed S shape; more specifically, a cumulated lognormal distribution

Figure 2.6.1. The Quasi-Engel curve at time t.



The model can also be used to generate predictions as to the shape of the aggregate diffusion curve over time. In order to do this, however, the time paths for μ_t, σ_t^2, μ_t^* and σ_t^{*2} must be specified. The simplest and most common assumptions are that:

$$\begin{aligned}
\mu_t &= \mu_o + g_1 t \\
\sigma_t^2 &= \sigma_o^2 \quad \text{for all } t \\
\mu_t^* &= \mu_o^* - g_2 t \\
\sigma_t^{*2} &= \sigma_o^{*2} \text{ for all } t
\end{aligned} \qquad (2.6.5.)$$

That is, constant growth of all incomes at the rate g_1, with the inequality of distribution remaining unchanged and a constant rate of decline in all critical incomes, with each individual's attitude remaining the same relative to his peers. The first pair of assumptions are fairly uncontroversial, the second pair amount to an overall change in tastes towards the new product over time due to bandwagon effects, greater information about the product etc. Perhaps one might argue with the exact form of the time path of y_{it}^* and the assumption that the growth in desire for the new product is uniform across the population, i.e. $\sigma_t^{*2} = \sigma_o^{*2}$; these assumptions may be modified, however, without destroying the simplicity of the approach.

One implication of these assumptions is that the Quasi-Engel curve shifts to the left over time. However, rather more interesting is the implied shape of the aggregate diffusion growth curve. This is derived by aggregating the Engel curve across the income distribution for each point in time. At time t, the expected number of owners is given by the sum of the probabilities at each income level, weighted by the number of individuals having that income. Algebraically:

$$m_t = n \int_0^{\infty} \Lambda(y_t | \mu_t^*, \sigma^{*2}) \quad d\Lambda\ (y_t | \ \mu_t, \sigma^2) \qquad (2.6.6.)$$

It can be shown [1] that by employing assumptions (2.6.5.), this aggregation leads to

$$m_t / n_t = N(\quad z_t | \ 0, 1) \qquad (2.6.7)$$

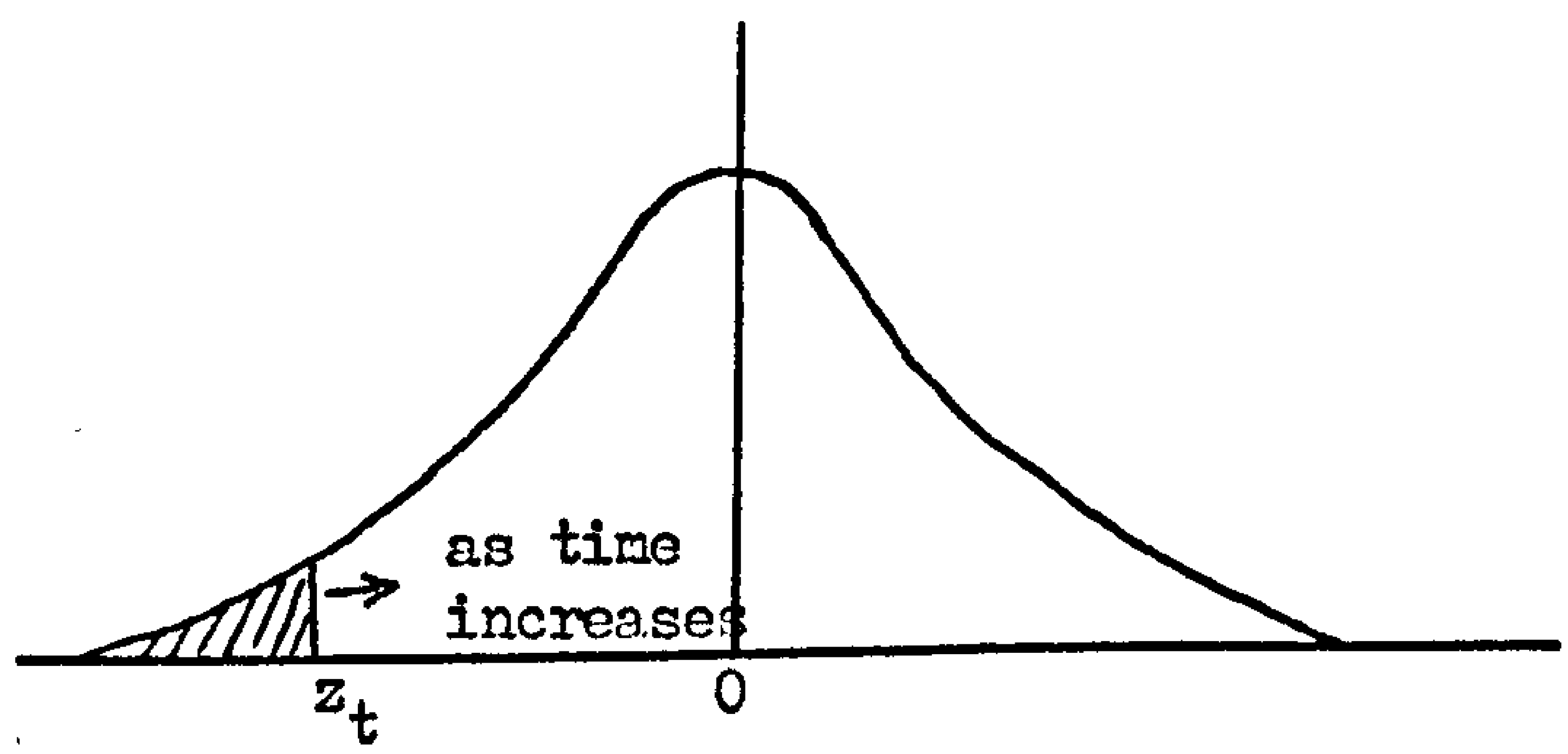
1. The formal, intermediate, steps are not presented here as the same sort of algebra appears in chapter 5. Alternatively, see Cramer, (1969, op.cit.) pp. 36 - 37.

$$\text{where} \quad z_t = (\mu_o - \mu_o^*) \; . \; (\sigma^{*2} + \sigma^2)^{-\frac{1}{2}} + (\varepsilon_1 + \varepsilon_2).(\sigma^{*2} + \sigma^2)^{-\frac{1}{2}}.t.$$

Thus, the expected diffusion at time t is equal to the proportion of a standard normal distribution to the left of z_t, where z_t is an increasing linear transform of time.

Graphically:

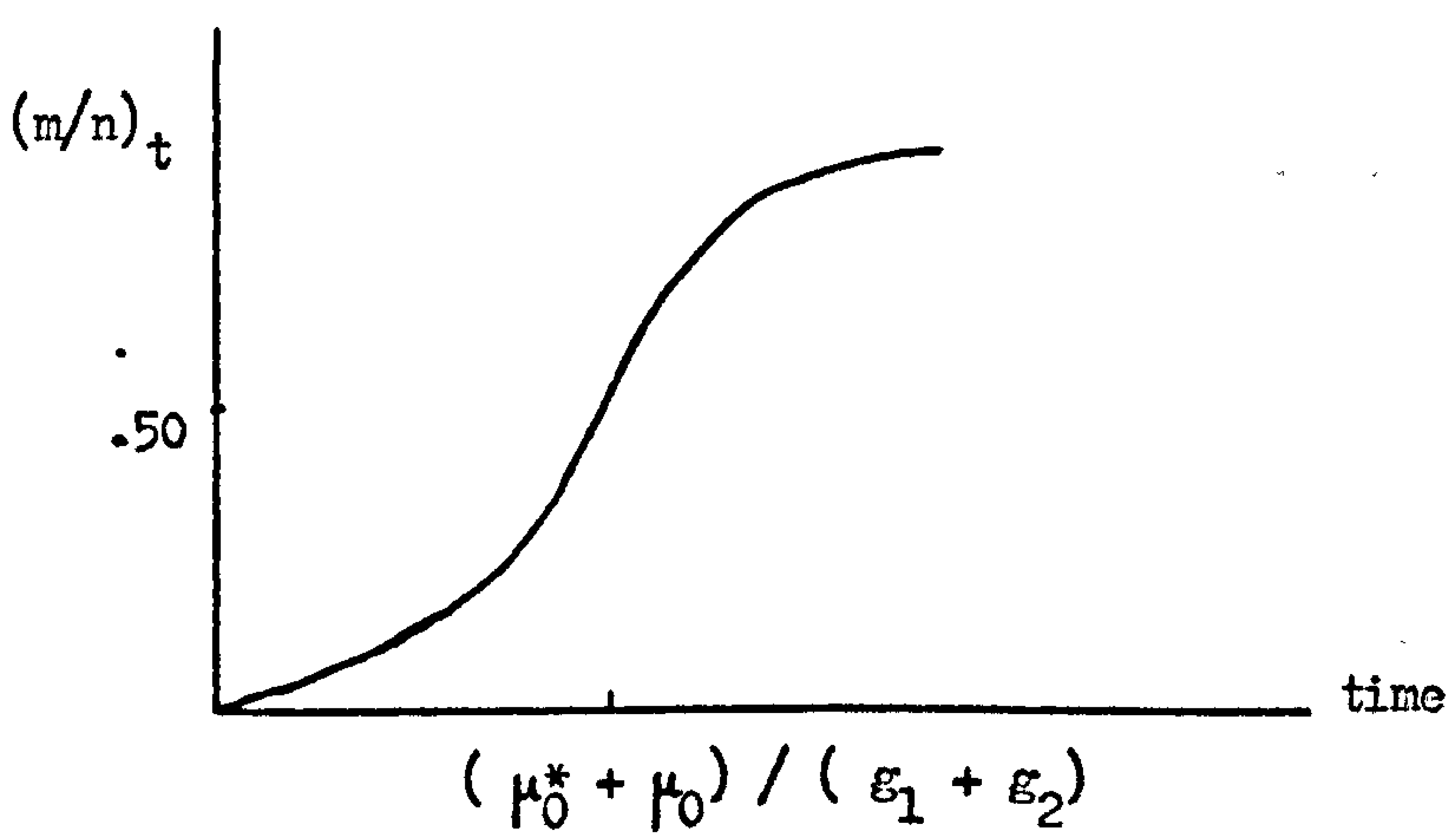
Figure 2.6.2. The standard normal.



$(m/n)_t$ is given by the shaded area as a proportion of the unitary total area under the curve. Moreover, z_t shifts to the right over time at a uniform speed.

Thus the growth curve predicted is a cumulative normal distribution, i.e. a symmetrical S shape which is nearly indistinguishable, in practice, from a logistic curve.

Figure 2.6.3. The cumulative normal curve.



A brief comparison with Mansfield's epidemic model is quite interesting. First, the shape of the growth curve is almost identical, but different assumptions about the time paths of y_{it} and y_{it}^* would modify this similarity. Second, a major theoretical improvement lies in the definite, if simple, hypothesis of the decision making at the individual level. Consequently, the choice of determinants of the speed of diffusion is not arbitrary but, instead, follows on as a precise result of this hypothesis, e.g. the growth rate of income and the mean and standard deviations of the income distribution.

In chapter 5, this basic model will be modified and extended to describe the diffusion of industrial processes.

7. A vintage approach to decision making.

It is quite obvious that very little of the research surveyed so far has used conventional economic theories of decision making. The lack of interest in the neo-classical theories is, perhaps, understandable; concepts of perfectly malleable capital and profit-maximizing seem a little inappropriate to an area characterised by uncertainty, imperfect information and possibly embodied technical progress. However, modifications to conventional theory are available which could render a potentially interesting analysis of diffusion. Specifically, the vintage model developed by Salter[1] and others would seem to offer a useful basis for such an analysis.

As a postscript to this survey, therefore, the implications of the vintage model are sketched very briefly.

New technologies are assumed to be embodied in new capital equipment and so gross investment is the vehicle of diffusion. Old equipment is only scrapped when its operating costs exceed the revenue it earns; new equipment is only installed if its total costs are covered by its revenue earning. This is, of course, the vintage model equivalent of marginal cost pricing, thus perfect competition and cost-minimizing are also assumed. Particularly important are the technological assumptions of indivisible plant, and that new technology cannot be introduced on old equipment. Thus, each vintage of equipment embodies the best practice technology of its date of construction, and is then committed to this technology (and a fixed labour complement) until its date of expiry.

The appearance of a cost-saving process innovation, within this framework, will have the following consequences. It becomes profitable to immediately replace some proportion of existing equipment because the total costs of the new equipment (including a profit allowance yielding the normal rate of return)

1. W. Salter, "Productivity and technical change," Cambridge University Press, 1960.

are lower than the operating costs of some old equipment.[1]

At the same time, product price will fall to equality with total costs of the new innovation - having previously been equal to the (higher) total costs of the latest vintage of the old technology. This fall in price is effected by the creation of new capacity, chasing the spectre of super-normal profits to be earned from installing the new innovation.

Some of the more efficient old technology equipment will still remain: its operating costs being lower than total costs of the new innovation. This equipment will gradually be replaced over the years, given a favourable movement in factor prices and vintage to vintage improvements in the new innovation. The important point is that 100% diffusion will not be instantaneous, even given perfect information and profit maximizing.

The above argument may be formalised mathematically fairly easily. Depending on the exact assumptions made about ex-ante substitutability, industrial structure, the nature of post-invention improvements in the technology (e.g. whether or not they are Harrod-neutral) etc., it can be shown that the diffusion growth path will be influenced by variables such as the growth in wages relative to prices, the growth of innovation price and the variance in the age of existing equipment etc. It would appear, however, that the assumptions of cost minimization, perfect information, indivisible plant and the embodiment hypothesis may only be relaxed at the cost of a considerable increase in algebraic complexity.

In a recent paper analysing the mechanization of reaping in the U.S. in the nineteenth century, David[2] similarly, attributes a central role to the movement over time in factor prices. Whilst it would be misleading to identify his model as vintage, it is based on similar behavioural assumptions. The basis of his explanation of the slow diffusion of mechanical reapers is that for many years, a majority of farms did not own sufficient acreage for the

1. The existing capital stock comprises a number of vintages embodying the old technology - the newer vintages being slightly more efficient than the old, due to minor improvements in specification over the years.
2. P.A.David, "Technical choice, innovation and economic growth," Cambridge University Press, 1975, chapter 4.

possible labour savings to pay for the cost of the new innovation. Diffusion only proceeded as the price of reapers fell relative to wage rates and as farm sizes increased.

His model may be summarised briefly as follows:

Adoption is profitable where $w_t\ (L_{iOt} - L_{iNt}) \geq P_{INt}$ (2.7.1.)

where w_t is the wage rate and P_{INt} the average annual cost of a mechanical reaper, L_{iOt} and L_{iNt} being the annual labour requirements of the old manual method and of the new mechanical reapers respectively for firm i, all at time t.

Assuming no scale economies in either method, $L_{iOt} = a_1 S_{it}$ (2.7.2.)

and $L_{iNt} = a_2 S_{it}$, where $a_1 > a_2$ (2.7.3.)

where S_{it} is the size of farm i at time t.

Combining (2.7.1.),(2.7.2.) and (2.7.3.) yields the condition for profitable adoption:

$$S_{it} \geq \frac{P_{INt}}{W_t} \left(\frac{1}{a_1 - a_2} \right) \qquad (2.7.4.)$$

Thus, diffusion only proceeds as a) farms increase in size, b) the cost of reapers relative to wage rates declines over time and c) reapers are technologically improved.

Whilst David does not have sufficient data to test these predictions rigorously, he does provide some evidence which indicates that (a) and (b) may have been important factors over the period considered.

As already mentioned, the vintage approach might be developed along similar lines and one would expect (b) and (c) above to figure centrally in any such development. The role attributed to firm size is particularly interesting: it derives <u>not</u> from economies of scale in the usual sense, but rather from the total lumpiness of the innovation - all farms must pay the same rental. Under such circumstances, large farms must always be at an advantage since the innovation price may be spread over a larger scale of operations. It is, perhaps, unlikely that this is the case for most 20th century industrial innovations and if capital costs are also proportional to the scale of operations, then the predictive power of this model largely disappears. Nevertheless, both David and Salter have provided a potentially richer framework for analysis, if only because of their concern with the individual firm's adoption decisions.

Clearly, it is necessary to examine the applicability of the technological assumptions of both models to innovations in the present sample before their value in this context may be assessed.

Conclusions.

It was stated in the introduction to this chapter that one of its main aims was to summarise the state of our knowledge about diffusion with respect to new process innovations. Sections 2 to 5 suggest that our knowledge is still very limited and tentative. It can be claimed, unequivocally, that the shape of the typical diffusion growth curve is sigmoid, but whether it is symmetrical or positively skewed is not certain. Mansfield, Griliches and Metcalfe all produce evidence to suggest that a major determinant of the slope of the sigmoid is the profitability of the innovation concerned. But there is no conclusive evidence on the role of industry-level characteristics in determining the speed of diffusion. Furthermore, very little is known of the reasons why some firms adopt innovations quicker than others. Two particular problems plague the empirical work in this area. Conventional firm-level characteristics appear not to possess much explanatory power, suggesting, perhaps, that non-quantifiable, technical factors play an important role at the individual firm level. Second it is relatively difficult to specify a dependent variable which will not produce biassed estimates of coefficients; some of the research in this area has used statistically imperfect measures and this, as much as technical or random factors, may have produced the poor results. The one conclusion that might be tentatively drawn from this work is that the speed with which a firm adopts a new innovation is directly related to its size. Finally, international differences in speed of diffusion can be partially ascribed to technological and institutional differences between the countries.

The second major aim of this survey was to search for a theoretical model on which to base the empirical analysis to be attempted later in the study. Unfortunately, such a basis is not apparent in the past work on diffusion of new process innovations. The epidemic model, used so regularly, suffers from two drawbacks. It depends on certain assumptions (often not made explicit) which are unlikely to apply in this context. Equally important, it produces an analysis that fails to recognize the individual firm's adoption decision as central to the aggregate diffusion process. Consequently, firms are viewed as an homogeneous mass thus pushing aside many of the more interesting theoretical problems. Even in the literature addressed specifically to explaining inter-firm differences, no attempt is made to formulate a fully-worked theory of firm decision-making. The work on international differences makes only a very limited theoretical contribution; on the other hand, it does indicate the importance of incorporating technological aspects into model building, even in non-international studies. Stoneman's use of the stock adjustment model does incorporate a definite theory of decision making but the overall framework is still inflexible and somewhat ad-hoc.

Perhaps the two most promising bases for theoretical development are to be found outside of the immediately relevant literature. The probit model, used in the analysis of the diffusion of consumer durables, whilst rather simplistic, does suggest a framework based on individual behaviour and which is flexible enough to be used in this area. Salter's vintage model also offers an avenue which, given certain modifications, might be used.

In the light of these conclusions, two major tasks become apparent. A fairly comprehensive survey is required of the technical characteristics of the sample innovations as listed in chapter 1. Specifically, the twin technological assumptions of the simple vintage model (embodiment of new technology and indivisibility of plant) should be tested. More generally, it is essential to establish whether new innovations have certain common attributes which might influence firms' decision making. Secondly, a realistic but empirically manageable approach to the firm-level adoption decision must be provided. The following two chapters are directed to these two tasks.

Chapter 3 : The Technical Characteristics of New Process Innovations.

As a preliminary to model building, this chapter presents an analysis of various technological aspects that seem to typify the sample innovations. It is, in many ways, a summary of Appendix one but also draws on past research findings where necessary. Section one considers, briefly, how representative the sample might be, in order that the conclusions of this chapter may be placed into some overall perspective of process innovations as a whole. Section two analyses the sources of the sample innovations; an important finding is that in nearly all cases, the innovations have originated from outside of the industries in which they are diffusing. Section three summarizes the main economic advantages and functions of the innovations. A crucial finding is that the technological assumptions of the simple vintage model do not fit a large proportion of the sample innovations. Section four assesses whether the innovations have fixed or variable coefficients of production. Sections five to eight investigate various aspects of the profitability of the innovations. Typically, profitability is found to vary over time and space for four main reasons. Across firms, differences in product, existing processes and inputs (Section five), all influence the returns from adoption; also important is the existence of scale economies in adoption (Section six). Over time, profitability often increases due to post-invention improvements in the technologies (section seven) and varies cyclically with the business cycle (Section eight.)

1. The sample innovations.

As can be seen from the Appendix, the industries in which these innovations are being adopted are distributed widely across the industrial spectrum: three process industries, five engineering, two textiles and three others. On the other hand, within these broad groups[1] the Iron and Steel, Weaving and Paper and Board industries between them account for thirteen of the twenty two innovations. The innovations themselves vary from some whose cost is measured

1. See Appendix one for an explanation of these industry classifications(e.g. table A.1.1.)

in hundreds of pounds (e.g. EH, ADH, SF[1]) to others costing millions of pounds (e.g. BOP, CC, ATL); potential profitability also varies widely; SF, for instance, having a typical payback of less than one year whilst SPC might typically need about eleven years to pay back.

It might be claimed, therefore, that the coverage effected is fairly broad[2]; whether it is representative of all process innovations must remain unclear, as very little is known about the characteristics of the typical process innovation. As is pointed out in the Appendix, the selection of the sample was non-random and so it is necessary to preface this chapter with the rider that the technical attributes which emerge do not necessarily apply to most process innovations. On the other hand, most of these attributes are common enough within the sample, and can be confirmed by reference to previous research, to suggest that their significance extends beyond this sample.

2. The sources of Invention.

No attempt will be made to distinguish the proportion of the sample innovations which are based on inventions made by individuals, as opposed to large industrial laboratories. As a brief study of Appendix 1 (or the technical references upon which it is based) will indicate, for many new technologies it is impossible to select one crucial invention. (Indeed, in a number of cases, early inventions were made last century e.g. TK, SL, CC, even though diffusion of all twenty two is essentially a post World War II phenomenon.) Moreover, the development of most of the innovations has been international and often the result of the coming together of ideas originating from totally different and independent sources. Designating a main source country for each innovation, it could be claimed that the U.K. was the originator or co-originator of six (but three of these are the minor

1. Virtually always in the main text, innovations will be referred to by their initials. Many of them have rather long and technical names and their repeated use would be rather irksome for both author and reader. A key is provided in table 1.1. of the introduction and table A.1.2. of Appendix 1.
2. See table A.1.1. of the Appendix.

weaving group); Germany and Austria of the four steel innovations, the U.S. of nine and Switzerland, Sweden and Canada each of at least one. This does not tell a very complete story however, U.K. firms played important roles in the development of at least twelve innovations; but then, this wider role also applies to the U.S. (more so) and Germany.

One conclusion of Pratten's[1], that is certainly confirmed by the sample, is the important role played by the capital goods industries and Research Associations. The Textile Research Association (better known as the 'Shirley Institute') was directly responsible for the three sizing innovations and the British Iron and Steel Research Association (B.I.S.R.A) was extremely active in the development of the steel innovations (particularly CC and VD.) Even more substantial was the role of the capital goods sector. In no case did the invention result from within the industry for which it was intended. Even more surprising is the fact, that, with the exceptions of ATL, the steel innovations and perhaps TC, the consuming industry played a totally passive role and did not contribute to the development of the innovation.

The conventional view of a firm inventing a new process for its own use, perhaps patenting it and then letting its competitors use the process under licence, is therefore totally inaccurate for the sample innovations. Rather, typically, the innovation is supplied by a capital goods industry, often under more than one brand name, with patents being less important than might be expected (perhaps because most of the scientific knowledge, on which the innovation is based, is usually relatively old and freely available.)

Clearly, the nature of the capital good industry and its relationship with the consuming industry may well influence diffusion. This will be considered in depth in the following chapter.

3. The functions and economic advantages of the innovations.

The innovations may be usefully grouped into four types. Six are essentially supplementary - in each case they are introduced alongside

1. C.F. Pratten, "Economies of scale in manufacturing industry," University of Cambridge Department of Applied Economics, Occasional papers: No.28, Cambridge University Press, 1971, p.293.

existing processes in order to speed the latter up. Usually, no existing technology is replaced. Six other innovations automate old manual operations or sometimes do away with them altogether. Again, there is not normally an old technology embodied in existing equipment which is to be replaced. Two others serve a different purpose from any existing equipment; for example, VD improves the quality of the product without replacing any old process. Only eight of the twenty two innovations conform to the technological assumptions of the basic vintage model. They embody the technique in new capital equipment and supercede an old technology also embodied in (existing) equipment. However, even in these cases, one of the simplifying vintage assumptions is violated: indivisibility of plant. For example, the Tunnel Kiln (TK) replaces an old technology embodied in Hoffmann Kilns, however a TK may be introduced without replacing the whole plant; the adjacent processes of clay-winning, shaping and drying do not need to be drastically altered and existing equipment may certainly be retained.

At this early stage, then, the simple vintage model can be ruled out. In the embryonic model set out in the last chapter, the sole reason for non-instantaneous diffusion was the efficiency of some existing equipment; for the fourteen non-vintage innovations in the sample, this is clearly inappropriate. The divisibility of plant in the remaining eight cases removes the analytical simplicity of the vintage model. No longer can product price be equated with the total costs of the latest vintage of the new innovation and the operating costs of the marginal equipment using the old technology. It is still possible to define the conditions for rational replacement of the old by the new equipment, but there is no longer the mechanical near-certainty that entrepreneurs will follow these conditions.[1]

1. In both cases of divisible or indivisible plant, rationality implies profit maximization. However, in the indivisible case, non-rationality leads to loss making on marginal plant as opposed to only a reduction in profits on the plant including the marginal equipment in the divisible case. Thus, it is much easier for entrepreneurs to 'carry' obsolete equipment when it is only part of the total process - the process, as a whole, still continuing to yield profits. See Salter,(1960) op.cit. pp.64 - 65.

No doubt the model could be adapted by introducing imperfect information and uncertainty into the analysis and introducing a more complicated pricing equation, but then the analytical simplicity of the model is lost.[1]

Returning to the sample innovations, it is possible to generalize about their main economic advantages. Labour saving is by far the most important attribute of nearly all the innovations. By definition this is true for the automation group, further the vintage group all have higher labour productivity than the old technologies. The main advantage of the supplementary group is in the speeding up of existing processes which, in turn, leads to more efficient use of labour (as well as other inputs.) Fuel savings are less pronounced but still significant for most of the vintage group and the supplementary group (for the reason just mentioned.) It is difficult to be precise about the extent of capital savings as this is rather an elusive concept in practice. Certainly, however, the supplementary innovations do increase the productivity of existing machines without any increase in other inputs. On the other hand, the vintage group, with one exception, do have higher investment costs per unit of capacity than the old technologies.

4. The coefficients of production of the innovations.

An area of some controversy in the theory of technical progress and formalized growth theory has centred on the coefficients of production. Are they constant or variable? In other words, is the production function a smooth curve or merely a single point? This controversy has never really permeated through to empirical work. However, it is a matter of some importance in this particular study. If new process innovations imply fixed capital - and labour-output ratios, then a given capacity using the new innovation can only be achieved with a given investment outlay: there will be no scope for firms to reduce their initial outlay by investing in a more labour-intensive version of the innovation.

1. Similarly, these technological points might be better analysed within an activity analysis framework but only at the loss of simplicity.

In fact it proves rather difficult to assess whether the sample innovations are characterised by fixed or variable coefficients. Data has been collected in some cases from which underlying production functions can be estimated. However, these are invariably ex-ante 'engineering functions' based on estimates of manufacturers or engineers. Typically, they do suggest fixed labour intensity, <u>for any given level of capacity</u>. In fact, evidence presented in section 6 suggests that a number of the innovations exhibit economies of scale in both labour and investment costs. Therefore both the labour- and capital-output ratios may be decreased by increasing the scale of the installation. However, once the scale is decided, both capital and labour inputs are not, generally, variable.

Nevertheless, given that these estimates are engineering or sales estimates, it is possible that they are geared to the prevailing wage rental ratio and it is not certain that under different factor price regimes, the manufacturers might not feel able to quote different specifications for their machines. A concrete example from the Appendix might clarify this point: a manufacturer of tunnel kilns reports a unique labour output ratio for a kiln capacity of ten million bricks p.a. which implies a certain fuel cost and labour cost per 1,000 bricks. In the event of, say, a large rise in the price of fuel relative to labour, it might be possible (perhaps by reducing automation at some stage) for the manufacturer to produce a similar sized kiln which is more labour intensive yet fuel saving. Such possibilities tend to be glossed over when taking engineering estimates too seriously.

This potential pitfall apart, it would seem that those innovations in the process industries (TK, BOP, CC, VD, VM) do imply basically fixed factor ratios, except to the extent that scale economies are not necessarily the same for different factors. Given the nature of the technologies (highly complicated with various sub-processes being highly interdependent and the sum total often requiring only two or three operatives,) it seems unlikely that there is much room for manipulation of factor ratios. Probably, choice is limited to choosing between slightly different alternative versions of the technology offered by different manufacturers.

A similar picture emerges for the innovations in the engineering industries (ATL, NCPP, NCTURN, NCTURB, SPC); although, for ATL at least, there is some choice in the extent of automation and for the other four there are a number of different brands on the market, each with slightly differing specifications. (However, most of these differences are due to variation in the tasks for which the machines are applied.)

The picture is a little hazy for the so-called supplementary innovations; four in the paper industry (SP, SF, F and WSB) consist simply of devices which are added to each paper machine and at this level, the coefficients are fixed quite clearly.[1] So too for Process control by computer of the paper machine (PCBC) although, again, the scale of the installation (dictated by the size of the paper machine) will determine the labour intensity of the installation. Computer typesetting (CT), although it involves substitution for a series of manual operations, is also characterised by fixed coefficients but the overall labour intensity of typesetting may be varied with the number of tasks which are computerised (i.e. the scale of the installation.)

Of the five innovations in the textile industries, ASB, EH and ADH are all standard pieces of equipment with unique labour intensities, as are SL; for the latter however, there is a certain amount of flexibility which derives from the type of SL chosen (Sulzer looms, for instance, being more capital intensive than the others,) and from the number of looms allocated to each operative. With tufted carpet machines, on the other hand, the only way of, say, reducing the labour intensity is by varying the size of the machine installed - but,as for the process innovations, different brands allow slightly different coefficients of production.

A tentative overall conclusion might be, therefore, that all of these

1. The same applies to GA in malting.

innovations involve basically fixed coefficients of production once the scale of installation (in terms of rated capacity) has been chosen. Certainly the continuous best practice production function of neo-classical economics seems inappropriate. Consequently, the initial outlay (dictated by capital intensity) is largely fixed for any given level of capacity.

5. Variability in profitability between firms due to technical factors.

As a brief reading of Appendix one will confirm, every innovation in the sample possesses certain technical characteristics which lead to sometimes quite large differentials in its profitability for different potential adopters.

Probably the most important contributor to this variability is the nature of the product which is to be produced using the innovation. In most industries, one can usually point to a handful of main products (and sometimes, even only one.) However using a finer definition of 'product', to include real or imagined brand differences, there are often as many products as there are firms. Very often the attributes which differentiate products, in the same industry, influence the returns the producer can derive from adoption of the new process innovation. Usually, the more differentiated and less standard the product, the less profitable is the new process. For instance, certain quality type bricks are not easily produced under the automation of TK; similarly, it is more difficult to produce carpets with a complicated design using TC; the continuous nature of CC makes it less beneficial for heterogeneous product mixes; GA is said to impair the quality of certain beers if applied in large quantities, and so on. Probably for just over half of the sample, it is the specialist products for which potential returns from adoption are reduced (but rarely removed altogether.) However, certain other innovations have the opposite effect: using NC, VM, VD and SPC, the quality and precision of the product can be improved considerably. Thus, although these innovations may be used for a wide variety of products, within the industry concerned, they are most beneficial for those products where quality is important.

In many cases, the nature of a firm's existing processes will also affect the returns to be gained from adoption. This is most obvious for the so-called 'vintage-model' type innovations. Clearly, the age and vintage of the capital equipment they replace will determine the returns from adoption. In the simple vintage model, of course, no firm ever replaces old equipment until it ceases to earn a surplus. But because, in practice (see section 3), plant is rarely ever indivisible, it is usually impossible to allocate to any sub-process the rent which it earns. Under such circumstances, the replacement decision is not automatic, rather, there is much more likelihood that firms will differ in their attitudes to replacement.[1]

On the other hand, certain of the 'supplementary' innovations operate more efficiently on relatively new equipment, e.g. GA, SP, SF, F, WSB. In these cases, therefore, adoption will be more profitable, the younger the firm's existing capital equipment. Similarly, SL are easier to accomodate in newer, better laid-out sheds,[2]

The technology embodied in adjacent processes is also sometimes important. For instance, VD is more compatible with the Basic Oxygen Steelmaking process than with the old-fashioned Open-Hearth; similarly, the use of GA only yields large savings if the kiln used for drying can handle the higher output-rate.

A third main cause of variability lies in differences between firms in the nature of their inputs, particularly raw materials. BOP can only be used when scrap metal constitutes no more than a limited proportion of raw materials; GA depends, for its profitability, on the type of barley used; TK are much easier to run when the clay used has a low carbon content; EH is less suitable for rayon and nylon and both SL and TC are less suitable for certain types of textiles.

1. Unless one is prepared to assume profit maximization and perfect information, both of which seem unlikely in these circumstances - these assumptions will be considered in more detail in the next chapter. Salter, himself, (op.cit.p.85) doubts the realism of such assumptions in the divisible plant case.

2. As Salter (op.cit. p.85) points out'a modern machine may have higher operating costs or a greater installation cost when installed in an older plant than the same machine in a completely modern plant.'

Fourth, and somewhat less tangible, the technical skills and educational attainments of managers, staff and labour force will also influence the returns to be gained from adoption. Many of the sample innovations are technically very sophisticated and can only be operated efficiently and balanced with existing plant if managers, staff and workers understand the technical complexities involved.[1]

Finally, in addition to these four general areas, a number of the innovations have certain peculiar attributes which will lead to differences, between potential adopters, in the returns to be had from adoption. For instance, Numerically controlled machine tools (NC) are most profitable if used to produce medium sized batches of output. Most engineering firms are limited in the extent to which they may vary their batch size, consequently, returns from adopting NC may vary, not only because of differences between firms in the areas already mentioned, but also because of different typical batch sizes.

Clearly, then, the potential profitability of adoption for any firm depends on so many technical factors that any empirical work will face quite serious measurement and specification problems.[2] Because this thesis will concentrate mostly on inter-industry differences, however, some of these problems can be circumvented. Without anticipating too much, what is fairly clear is that for any innovation diffusing in any industry, there will be a distribution, across firms, of the profitability of adoption. Whilst it is unlikely that an estimate of profitability for each firm can be obtained, there is, perhaps, more chance of discovering something about the mean and variance of the distribution, for each innovation in the sample. Indeed, estimates are available for the means in appendix one. Whilst estimates of the variance

1. Of course, expertise may be 'bought-in' at the time of adoption, but this in itself will increase the cost of the new process. Either way, firms with less skilled managers stand to gain less from adoption.
2. The rather poor explanatory power of past work on inter-firm differences is, therefore, hardly surprising (see section 3 of the previous chapter.)

are usually unavailable, the above discussion suggests a number of measurable variables which might help explain differences between industries and innovations in these variances of profitability.

For the moment, however, the empirical problems are postponed for later discussion. One interesting implication of this section which may be noted is that one of the basic assumptions of the logistic model appears to be unrealistic. Clearly, all firms are not equally susceptible to the new innovation.[1]

6. Economies of scale and other size advantages.

A possibility, largely ignored in most past work in the diffusion field,[2] is that the savings emanating from adoption of any innovation may well be determined by the scale on which it is adopted.

However, for the sample innovations, statements appear in the technical literature to the effect that the investment cost per unit of capacity decreases as the capacity of the installation increases. In some cases, data is available which suggests that a good algebraic approximation of this relationship is given by:

$$\frac{K}{S} = \alpha S^{\beta} \qquad (3.6.1.)$$

where $\frac{K}{S}$ is the investment cost per unit of capacity and S is capacity. α and β are constants.

In the appendix, (3.6.1.) has been fitted to data for 3 of the innovations, the estimates of β being as follows: -.453 (ATL); -.32 (TK); -.20 (BOP). For VD, there is also a close inverse relationship of the same sort between investment costs and ladle capacity, which, in turn, will help determine overall capacity.

1. Thus violating assumption (c) of the epidemic model as outlined in section one of the previous chapter.

2. But see a footnote in Mansfield (1968) p.160, which Mansfield then subsequently ignores. Furthermore, elsewhere he explicitly assumes that there are no scale economies: (Mansfield, 1963, op.cit., footnote p.292) See also David (1966) for a notable exception.

Leckie and Morris [1] have found that a similar form to (3.6.1.) describes scale economies in the investment costs of virtually all new processes in the Iron and Steel industry. Thus, it is likely that similar economies apply to the other two sample innovations in that industry (that is CC and VM.)

The conventional view is that computers, too, exhibit investment scale economies. This view has been challenged by Stoneman who presents econometrically estimated relationships between computer price and three aspects of computer size, which are not inconsistent with constant returns to scale. [2] Judgement is perhaps best reserved, therefore, for CT and PCBC- the two computer-using innovations in the sample.

For the fifteen other sample innovations, there appears to be no evidence either way on this matter. It is noticeable that the seven for which some evidence was available are all relatively expensive, occur mainly in the process industries and fit the vintage model assumptions quite well.

All of this should hardly come as a surprise; (3.6.1.) has often been used in past research as a description of scale economies in investment costs of new plant, particularly in the process industries. Haldi and Whitcomb[3] have estimated β for 687 items of equipment, used mainly in the process industries, on the basis of data collected predominantly from industrial catalogues. For only fifteen items was $\hat{\beta}$ positive, over 80% had $\hat{\beta} < -.2$ and over 70% of these recorded $-.6<\hat{\beta}<-.3$. Bruni[4] reports similar findings for oil and petro-chemical plants. His results yield a distribution for $\hat{\beta}$ with a mode between $-.3$ and $-.4$. This is broadly in agreement with the well known 'six-tenths' rule of engineering.

1. A. Leckie and A. Morris, 'The effect of plant and works scale on costs in the Iron and Steel Industry', Journal of the Iron and Steel Institute, May 1968 pp 442-452, in which they find $-.33<\beta<-.25$.
2. Stoneman, (1974) op.cit. - he also presents a survey of the conventional empirical findings of increased returns.

3. J. Haldi and D. Whitcomb, 'Economies of scale in industrial plants,' Journal of Political Economy, vol. 75, 1967.
4. L. Bruni, 'Internal economies of scale with a given technique.' Journal of Industrial Economics, June 1964.

It is usually argued that this rule applies to most equipment which consists of cylinders, spheres, tanks, tubes etc. In such cases (mostly process equipment) production capacity is determined essentially by volume, whilst capital cost depends more on the surface area of the vessels. Surface area, after all, dictates the quantities of materials and physical effort needed to construct the equipment. Basic mathematics shows, of course, that to increase volume, surface area needs to be increased by smaller proportions.[1] Using the notation of (3.6.1.), the rule can be expressed as

$$K = \alpha S^{.6} \qquad (3.6.2.)$$

which is in line with a value of β of -.4.

Whether such economies have an upper limit is not certain; Bruni suggests that at certain very high capacities, stresses in the raw materials may appear that cannot be supported. The data which is available for the sample innovations does suggest that there are upper limits for BOP, ATL, and TK, but that these upper limits appear to be at capacities in excess of anything ever installed in the U.K.

Finally it should be remembered that virtually all of the above evidence is based on ex-ante engineering data and not actual installations.[2] These, then, are the theoretical scale economies, not the actual, achieved economies.

Engineering data is also available on scale economies in the operating costs of some of the sample innovations. A log-linear form again appears to be a reasonable approximation:

$$\frac{OC}{S} = a\,S^{b} \qquad (3.6.3.)$$

where OC = operating costs, S capacity, and a and b are constants.

The following estimates of b were obtained: -.21(ATL); -.13(TK); -.3(BOP) and -.52(VD). That is, slightly smaller than the investment economies for ATL and TK but slightly larger for BOP.[3]

1. When comparing two similarly shaped solids, the ratio of surface areas equals the ratio of their volumes to the power 2/3.
2. One implicit assumption, for instance, being that large installations do not take longer to construct. If, in reality, they do, then higher interest costs will reduce the scale economies.
3. It should be stressed that all these estimates are based on only a handful of observations and are, as such, very rough.

Leckie and Morris[1] used a slightly different functional form[2] to describe operating cost economies of Iron and Steel processes:

$$\frac{OC}{S} = \frac{b_1 q + b_2 S^{b_3}}{S} \qquad (3.6.4.)$$

where q = actual output and b_1, and b_2 and b_3 are constants. This formulation differentiates between those operating costs which are relatively fixed (and thus related to capacity), and those which are more variable (and thus related to actual output.)

In fact, their estimates of b_3 fell into the range .67 to .75 which implies a range for b in (3.6.3.) of -.25 to -.33. It is likely, therefore, that had data been available in this study for VM and CC, similar economies would have emerged as for ATL, TK, BOP and VD.

For CT and PCBC, once more, there is no direct data, but Stoneman's analysis, this time, does suggest economies in operating costs for computers as a whole. Furthermore, it is known that the quality of service offered by computers for PCBC and CT improves as the scale of the installation increases. This is largely due to increasing scope for automation and thus labour savings.

Again, there is considerable evidence, in past studies, of substantial economies of scale in operating costs for process equipment. Haldi and Whitcomb[3] estimated (3.6.3.) for a sample of 32 processes. In all cases, $\hat{b} < 0$ and in 19 $-.3 < \hat{b} < -.1$. If only labour costs are considered, the economies are even more pronounced. Including some earlier results of Isard and Schooler[4] in their sample, they conclude that 37 of 52 estimates of b are smaller than -.6, when OC = labour costs. Bruni's[5] work suggests an average value for $\hat{b}$ of -.76 when only labour costs are considered.

Similar substantial economies in labour usage have also been found using British data. Again for innovations in the process industries, Pratten[6]

1. (1968) op.cit., on the basis of engineering data.
2. Which differs from the approximation of (3.6.3.) only in the inclusion of an additive constant.
3. op.cit.
4. W. Isard and E. Schooler, 'Location factors in the petro-chemicals industry.' U.S. Department of Commerce, 1955.
5. op.cit.
6. op.cit., p.12.

finds that in many cases, total labour requirements were relatively insensitive to the scale of installation and maintenance costs were largely proportional to investment costs (which, as has been suggested, exhibit scale economies.) He also doubts[1] whether there will be any significant dis-economies of management for individual plants, as the control problem is not notably increased. Yet again, however, the deficiencies of engineering estimates must be acknowledged. For instance, it is possible that factor prices will be sensitive to the scale at which the factors are employed - capital costs will quite possibly rise with scale, so too might wage rates.

As far as is known, there is no engineering evidence of scale economies for the 15 other sample innovations. However, the technical sources considered in the Appendix suggest that, for a number of the sample, there are so-called Economies of large numbers. For example, although unit costs of shuttleless looms (SL) may not be sensitive to the size of the loom, they are to the number of looms installed, (typically, looms and some of the other innovations are installed in batches.) For SL, TC, NC, SPC, CT and PCBC, such a numbers effect may be attributed to (a) economies in servicing and programming, (b) a proportional reduction in necessary stocks of spare parts,and probably most important, (c) a proportional reduction in 'setting-up' time. Another effect peculiar to SL and TC derives because labour is 'lumpy' in some processes in textiles, that is, one operative is responsible for a number of looms. Thus, the larger the number of looms installed, the better chance there is of optimizing the capital-labour ratio..

It seems safe to conclude, therefore, that significant economies of scale in operating and / or investment costs obtain for virtually all of the sample innovations. They are, perhaps, most pronounced for the more technically sophisticated innovations which are often to be used in the process industries.

Whether these produce greater incentives for large firms depends, very much, on the extent to which a firm's size limits the scale at which it can adopt.

1. ibid., p.299.

For the 'lumpy' innovations, there is little or no flexibility; for example, a brickmaker with an output of 10 million bricks per annum will not be able to install a Tunnel Kiln (TK) with a capacity significantly different from that figure. Apart from local demand conditions, he will be constrained by the capacity of the adjacent processes such as clay winning, shaping and drying. In other words, the need for balance will often restrict the scale of adoption. However, in some industries large firms are not bigger because their individual plants are larger but, rather, because they have more plants. Under these circumstances, it is possible that smaller firms are just as likely to install large units of the innovation as are large firms. On the other hand, just because the large firm operates more plants, it has a better chance of having what might be termed the 'ideal conditions' for adoption.

Even should small firms be as able to adopt on much the same scale as large firms, they may still be unwilling to do so because of the proportionately greater inroads it might make on other objectives. For instance, it would be relatively more difficult for small firms to avoid the disruption of work-flow and loss of output during the change-over period[1] involved by larger scale adoption.

One possibility which might reduce the scale advantages would be the existence of scale economies in the old technology being replaced. This can be discounted immediately for the six innovations replacing manual operations and the six supplementary innovations. For the other ten, the possibility remains although with one exception, there is no evidence either way. For CC there is evidence that the old technology did also exhibit scale economies. Pratten provides some detail on this example in his comprehensive survey of scale economies in production processes.[2] However, he does concludes that "many new processes are increasing the economies of scale and increasing their range, and though some new techniques ...... reduce the economies of scale, the impression gained from the industry studies was that these are exceptional cases."[3]

1. It may also be more difficult for small firms to retain flexibility: many new innovations increase the standardization of the end-product.
2. Pratten, op.cit.
3. ibid. p.303

In addition, there are a number of <u>other technical attributes of new innovations which make adoption more favourable, at an early stage, for large firms</u>. First, large firms may be able to use new innovations more intensively: the importance of avoiding 'down-time' for CT and PCBC was stressed by manufacturers; similarly, large firms may be able to operate SL's for longer runs, not least because they are often able, through localised monopsonist power over labour, to demand shift working.

Second, large firms are more likely, perhaps, to employ the skilled management and staff needed to understand the technical intricacies of new innovations, especially: GA, NC, TK, PCBC, CT and SPC.

Third, most of the large innovations often require more than a year for proper installation. As suggested above, this can often lead to loss of output (where building on an old site) and problems of liquidity in the change-over period. Similarly, periods of retraining (and thus further loss of production) may be needed for managers, staff and workers. Because of their greater resources, larger firms may be better able to absorb these disruptive effects.

Fourth, and as mentioned briefly above, is the higher probability of large firms having the ideal conditions for adoption. As indicated in the previous section, although all of the sample innovations are suitable for most firms in the appropriate industries, they are more suitable for some than others, (in terms of having the most appropriate product/process/input etc.) "Large firms, because they encompass a wider range of operating conditions, have a better chance of containing those conditions..."[1] A second possibility, also noted by Mansfield, is that larger firms will have a greater probability just because of their size, of having a unit of the old technology needing replacement (perhaps failing to earn rent or nearing the end of its physical life.) At least, this is true, so long as there is no pronounced tendency for large firms to have newer capital equipment than small firms. McGee [2] postulates that this may not be the

1. E. Mansfield. (1968), op.cit. p.156.
2. J.S. McGee, "In Defence of Industrial Concentration." New York, 1971, pp. 113 - 115.

case. He argues that if large firms are large because they have been progressive in the past, then it is possible that their capital stock may be younger than the average. However, this point is only valid if the large firm has expanded in recent years, which need not always be the case. Indeed, it is possible that due to their market power, they are able to 'carry' old and inefficient equipment longer. Yet if this is the case, there seems no reason why they should decide to replace that equipment now. As Meyer and Kuh find[1], "firms which, on average, have older equipment tend to keep it that way."

An indirect example of the strength of this 'ideal conditions' argument is provided by Hakonson's work on the international diffusion of Special Presses.[2] For a number of countries, he divided potentially adopting paper firms into four groups (according to the speed at which they had adopted the innovation.) Having collected an enormous amount of technical information on the paper machines in all of his sample firms, he computed, for each of these machines, an expected pay-off period[3] from the adoption of the innovation on to that machine. Taking only the most profitable case (machine) for each firm, he cites average pay-out[4] and average size of firm for each of the four groups of firms. For each country, there is a quite obvious tendency for the groups with the largest average firm size to have the most profitable machines down to the groups with the smallest firm size having the least attractive machines on which to adopt SP.[5] As far as is known, SP has no economies of scale; thus this tendency

1. J.R.Meyer and E. Kuh, "The Investment Decision", Cambridge (Mass.), Harvard University Press, 1957. Although this finding is to be found in the above reference, the quote is from Meyer's chapter in "New Industrial Processes", op.cit. p.171.
2. S. Hakonson, op.cit. S.P. is a supplementary type innovation which is added to existing paper machines.
3. Based on technical factors such as age, speed and width of the paper machine, thus, this is an ex-ante concept.
4. ibid. p.77, table 4.7.
5. ibid. p.83, table 4.13. Strangely Hakonson appears not to have noticed this close relationship between profitability and firm size. A recognition of it might have helped him to interpret some of his later empirical results.

for large firms to have the most potentially profitable machines can only be attributed to the 'ideal conditions size advantage.' That is, large firms tend to find that their most profitable opportunity is more profitable than are the smaller firms' simply because the larger firms have more machines.

7. Post-invention improvements in new innovations.

Rarely, if ever, does the technological development of a new process end once it has been adopted for the first time. The manufacturers usually continue to divert resources to effect post-invention improvements for some time after the date of introduction.[1] The technical literature surveyed in Appendix one repeatedly comments on 'rapid technological developments' (or similar phrases) over a number of years. Whilst numerical evidence is often unavailable, these developments appear to exist in four main areas. Perhaps most quoted, are improvements in specifications leading to increased productivity of the variable inputs (mainly labour) when using the new innovation. Also fairly common, are reports of a steady decrease in the quality adjusted price of the innovations. Other improvements sometimes lead to increasing applicability to wider operating conditions for the new innovations and, sometimes, increases in the extent of scale economies.

Where numerical evidence is available, labour productivity, particularly, of later vintages of the innovation is often far higher than for the earlier vintages, (see, for instance, TK and TC within this sample.) Similarly, the data that is available suggests that manufacturers are able to reduce price over the first years, at least, of the life of some processes, e.g. GA and CT and computers generally.

Unfortunately, past work on diffusion has little to say, empirically or theoretically, on this matter.[2] Yet as early as 1958, Enos presented data

1. In certain extreme cases, these post-invention improvements may be so large as to raise the definitional problem of whether the same innovation is still being studied. One criterion for inclusion of innovations in the present sample was, in fact, that in spite of improvements they should exhibit broadly stable technological functions and designs over the diffusion period.

2. With the exceptions of David's contributions (op.cit.)

showing quite clearly that the investment costs of four innovations in petrol refining fell dramatically over a number of years.

More generally, the considerable evidence on "learning by doing"[1] is particularly relevant to this area. The arguments for learning by doing are, by now, well known: as manufacturers have more and more experience at producing a new product, they acquire a better understanding of the technical relationships involved and their workers acquire new on-the-job skills. Thus, the costs of producing new innovations might be expected to fall in real terms for a number of years after the initial introduction: "At one firm it was suggested that the costs for the initial batch produced may be as much as three times the average cost of a machine tool after it had been in production for eighteen months or so. Another 'rule of thumb' suggested was that the average cost for the first production batch could be reduced by more than one third."[2]

When statistical description is possible, the learning curve is usually estimated in one of two forms. Enos,[3] for instance, found that a good description for the four petrol refining innovations was provided by the curve:

$$K_t = A\,t^{-.45} \qquad (3.7.1.)$$

where K_t is per unit investment costs at time t (that is the unit price of the innovation), A, a constant and t the number of years elapsing since the innovation's first introduction.

Alternatively, Hirsch[4] used the form:

$$L_t = a\,Y_t^b \qquad (3.7.2.)$$

where L_t is labour requirements per unit of output at time t and Y_t the <u>cumulative</u> number of units of the new product produced by time t. His

1. Stimulated by K.J. Arrow's paper, "The economic implications of learning by doing." Review of Economic Studies, June 1962.
2. Pratten, op.cit., p.167.
3. J.L. Enos, "A measure of the rate of technological progress in the petroleum refining industry," Journal of Industrial Economics, June 1958 pp.180 - 197.
4. W.Z. Hirsch, "Firm Progress Ratios", Econometrica, April 1956, pp.136-143.

estimates of b varied between -.29 and -.41, the products concerned being 22 new capital goods (including machine tools, textile machinery, construction machines, airframes and ships.) In another study of airframes by Hartley,[1] an estimate of $b = -.32$ was implied. (It should be stressed that L_t is labour requirements of the producer and not the user of the innovation.)

These two alternative specifications are only exactly equivalent given certain assumptions about manufacturers' pricing, the growth path of wages and other variable costs.[2] As the whole question of manufacturers' pricing behaviour is discussed in more detail in the next chapter, at this stage, the only conclusion made is that manufacturers' costs of production, in real terms, are likely fall over time. Both (3.7.1.) and (3.7.2.) imply this.

Switching the emphasis to the consumer (potential adopter), this discussion is relevant to the price he pays for the innovation, but equally important is the movement over time in, say, the labour-output ratio associated with operating the new process.

A distinction is needed here, between improvements due to learning by the consumer and by the manufacturer. For there to be any symmetry in the argument, one would expect that, when a new innovation is installed, there will be a period in which the adopter gets to grips with its new technology: workers learn new skills, engineers make running improvements peculiar to

1. K. Hartley, "The learning curve and its application to the Aircraft industry," Journal of Industrial Economics, March 1965, pp.122 - 128.

2. For instance, if manufacturers are average cost pricers, wages grow according to the function $w_t = a_1 t^{b_1}$, cumulative output grows according to the function $Y_t = a_2 t^{b_2}$ and labour is the only cost, then

$$K_t = (1+M)\ w_t L_t = (1+M)\ a_1 t^{b_1} . a . a_2^{\ b}\ t^{b_2 b} . = (1+M)(a . a_1 a_2^{\ b})\ t^{b_1 + b . b_2}$$

(M being the constant mark-up rate.)

Alternatively, marginal cost pricing with constant returns at every point in time would ensure the same sort of result.

the local conditions etc. It would also seem likely that there is some feed-back (especially given that an after-sales service will continue the consumer-manufacturer dialogue,) to the manufacturer, who may then make adjustments to the underlying technology for future vintages. It is also likely that research and development work continues, even after the innovation has started to diffuse. This is certainly the case for many of the sample innovations; advertising material emphasises this aspect, especially where there appears to be competition between different brands of the same innovation.

It would not seem unlikely, therefore, to assume for L_{Nt} - labour input per unit of output needed using vintage t of the new innovation - a relationship to the year of the vintage t of the form:

$$L_{Nt} = A_2 t^{B_2} \qquad (3.7.3.)$$

where A_2 and B_2 are constants, $B_2 < 0$.

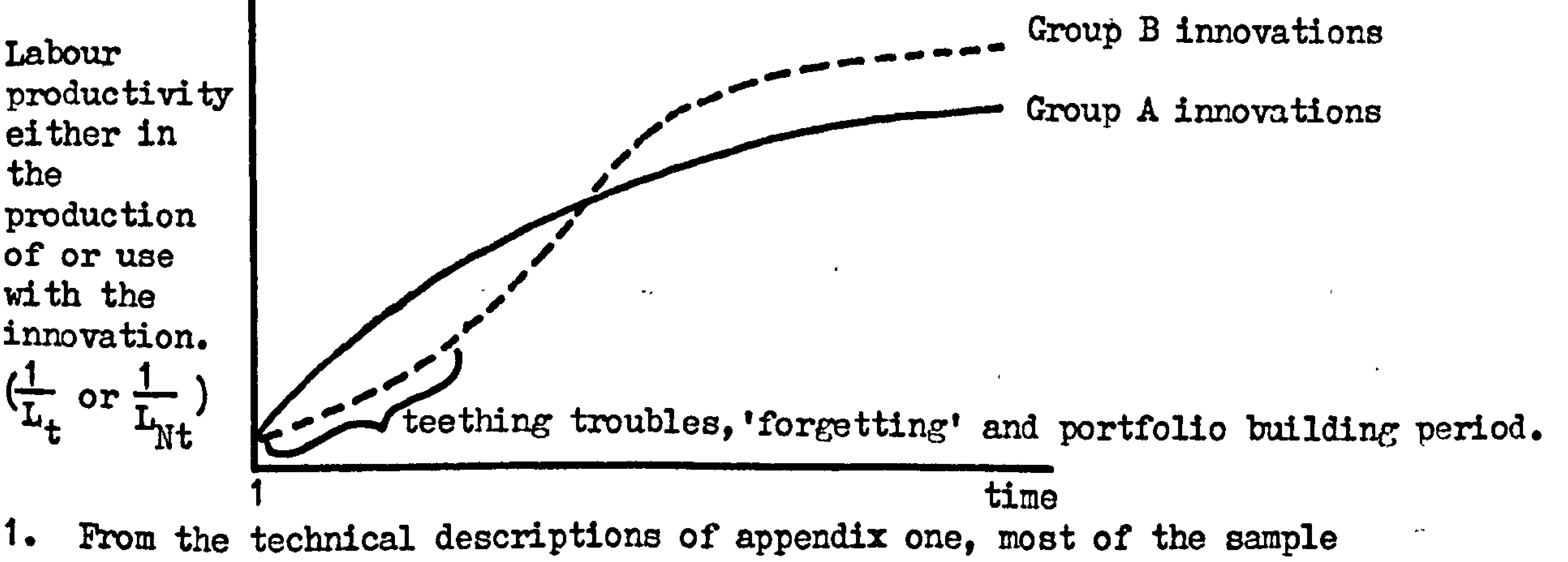
That is, the same sort of learning curve as above. Such a form is not inconsistent with the little data available from Appendix one and the frequent assertions, in the technical literature, of improvements from vintage to vintage in many of the innovations.

There are a number of other conclusions drawn in the above past work on learning which are often corroborated in Appendix one. The scope for learning appears to be greater (a) the more technically sophisticated the process innovation and (b) the more discontinuities or breaks in production there are (for the manufacturer.) Likewise, both factors tend to produce significant learning effects for a longer period. On the other hand, discontinuities in production tend to produce more 'forgetting', especially in the early years. Another tendency which is quite pronounced for some of the sample innovations (CC and VD particularly, but also most of the other sophisticated innovations) is that of 'teething troubles'. In other words, significant technological stumbling blocks appear both in the production and the use of the innovations in the early years. Similarly, some of the innovations are so lumpy that each potential adopter will only ever need one or two units of them. From the manufacturers' point of view, this means greater heterogeneity in their installations; this is especially so given the complexity of most of the lumpy innovations, which also tends to produce inter-site differences because of different operating conditions.

For these reasons, it is worth differentiating two types of innovation. Group A can be defined as technologically simple, probably relatively cheap and produced off-site. Group B are more sophisticated, expensive innovations which are produced on a one-off basis, often requiring lengthy periods of installation on the adopter's site.[1] Learning effects (as manifested in falling labour inputs, both in the production of the innovation and when using it) for group A might be initially quite large, but soon falling away drastically. For group B, they are likely to be much longer lived and, in the long run, more substantial. Nevertheless, over the early years, whilst the manufacturer is overcoming teething troubles and building up a portfolio of knowledge about his customers' different operating conditions, learning may be quite limited for group B.

These two alternative hypotheses may be stylized graphically as in figure 3.7.1.

figure 3.7.1. The learning curves for Group A and Group B innovations.



1. From the technical descriptions of appendix one, most of the sample innovations fit quite clearly into one or other of these categories. A should include SP, SF, F, WSB, GA, ADH, EH, ASB; B should include BOP, VD, VM, CC, ATL, PCBC, and TK; the other seven share some characteristics with both groups.

Thus, in the long run, both curves settle down to the linear in logs mathematical form with decreasing returns; in the short run, Group B might possess a learning curve more akin to a rough exponential. This distinction is important. For instance, Enos's data (for what were obviously group B innovations), covered many decades, and thus might be expected to take the conventional form. In this study, the time span is much shorter. In considering inter-firm diffusion, in which only the first installation of the innovation for each firm matters, only a small part of the life of the innovation is analysed. In the steel industry, for example, (but also many others), a technology may reign supreme for as long as a century, but it may be only two decades or less before all firms have made their first purchase of the technology. Consequently an exponential learning curve may be more appropriate for some innovations in this context.

The two other aspects of 'technological learning' mentioned earlier are well documented:

a) in many cases, the range of operating conditions for which the innovation is suitable is extended widely in the very early years. This would appear to have been true for at least F, SF, SL, SPC, TC, BOP, CC, VD and VM.

b) in a few cases, the economies of scale inherent in the technology have been increased in the early years (probably only for VD, BOP and CC within the sample.)

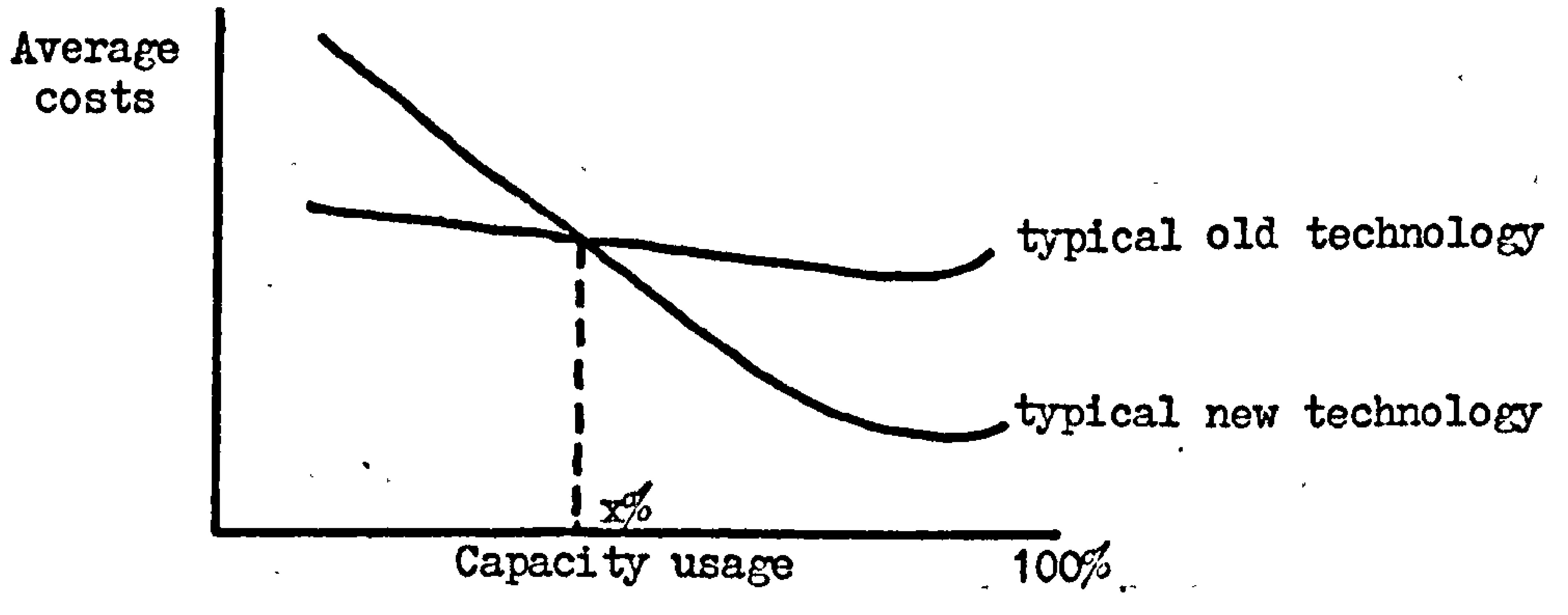
Both of these phenomona have been due, certainly, to feed-back from adopters. However, it should be stressed that most of these improvements were made before the innovation in question had diffused to any extent and, sometimes, even before it had been adopted at all in this country. Therefore, both cases will be largely ignored in subsequent model building, although it is interesting to note, in passing, that this might be considered an advantage in not being the innovating country.

Finally, one other aspect of learning which should be acknowledged (although, again, it is doubtful whether it is of general relevance to the sample innovations) is the so-called '<u>sailing ship effect</u>'.[1] This is said to occur when the introduction of a new innovation leads to a sudden spurt of improvements in the old technology; this happened to a certain extent for TK (abroad but not in this country) and BOP (also mainly abroad), another example is cited by Nabseth [2] in Sweden.

8. The short-run average cost curve of the new innovation and other cyclical considerations.

For a number of fairly obvious reasons, the attractiveness of adoption of a new innovation is likely to vary across the trade cycle. Over half of the sample innovations are typified by high fixed costs relative to the old technology. These are due, variously, to the need to employ highly paid and highly skilled staff and workers, high capital charges and high start-up costs. Although variable costs are lower, reasonably high output rates are still required for these to outweigh the fixed cost differential. This applies to the six continuous processes (usually replacing old batch-type processes) and the six innovations automating old manual operations.

Figure 3.8.1. Short run average cost curves.



In terms of figure 3.8.1. capacity usage would have to be higher than x% for there to be any savings from using the new innovation.

1. The name derives from the improvements in sailing ships that followed the first introduction of steam ships, see C. Freeman, "The economics of industrial innovation." Penguin, London 1975, p.47.
2. Nabseth, (1973), op.cit.

There is a further set of eight innovations which produce unit cost savings mainly by speeding up existing processes, (without an increase in most inputs), or by removing bottlenecks. For these too, therefore, adoption is only worthwhile given sufficient demand. Similarly for SL, most of the cost savings only result from long production runs which also depend on reasonably healthy demand conditions.

Clearly, then, the returns to be gained from adoption will vary according to conditions and expected capacity usage. This may be particularly important if firms' investment decisions are based on simple rules of thumb, such as the pay-back method, which weight immediate returns as overwhelmingly important.

On the other hand, as has been already mentioned, many of the innovations require lengthy building programmes which lead to short-run disruption and output losses - these may <u>encourage</u> adoption in periods of low demand when disruption and output losses may be less real.

The implications of the first argument are interesting. In industries with highly fluctuating demand, diffusion may proceed at a slower pace if entrepreneurs prefer to operate along a reasonably flat short-run average cost curve as for the typical old technology in the diagram. Moreover, large firms may be at an advantage. They may be more able to optimise their technology mix so as to operate the new innovation at full capacity permanently, and to satisfy the fluctuating residual by retaining some units of the old technology. Indivisibilities may prevent many small firms from following this option.

Conclusions.

It was suggested in the concluding remarks of the previous chapter that a detailed analysis of the technological characteristics of the sample innovations was necessary before embarking on economic model building. Certainly, a number of quite important findings have emerged from such an analysis. First, on the negative side, the technological assumptions of the simple vintage model do not appear appropriate for most of the sample innovations. Furthermore, at least two of the implicit assumptions of the Mansfield model of inter-firm diffusion seem to be invalid; the profitability of the typical innovation is constant neither across firms at one point in time, nor for the same firm over time.

On the positive side, a number of characteristics of new innovations have emerged which will be incorporated into the model of chapter 5. Typically, new innovations originate from capital goods industries who subsequently offer them for sale to the consuming industries. They are invariably mainly labour saving with fixed coefficients of production (both of these conclusions must be judged in light of rather poor quality data.) The profitability of the typical innovation varies across firms according to the technical characteristics of the innovation and the firm concerned. Due to a number of factors, including increasing returns to scale in the use of many innovations, profitability should be positively correlated with the size of the potential adopter. Furthermore, the profitability will vary over time due to capacity fluctuations and post-invention improvements in the underlying technology, the latter being mainly due to 'learning-by-doing'. It is postulated that the learning curve may take one of two different shapes over the diffusion period due to certain important technical differences between simple, class A innovations and the more sophisticated, lumpy class B innovations.

Finally, it should be stressed that <u>most</u> of the analysis rests on technical information (surveyed in appendix one) about the 22 sample innovations. Thus generality to all new process innovations cannot be claimed, although in many ways the sample is quite broad and perhaps reasonably representative.

Chapter 4 : The Adoption Decision.

The purpose of this chapter is to provide the economic basis of the model to be developed in chapter 5. Specifically, some statement is necessary on the nature of the firm's adoption decision with respect to new innovations.

It is possible, straight away, to rule out standard profit maximising as a description of the behaviour of entrepreneurs with respect to new process innovations. None of the sample innovations had diffused 100% in the relevant industries within the first ten years of their existence. Yet, in virtually all 22 cases, the average pay-back associated with their adoption was in excess of the typical rate aimed for by industry on capital investment, generally.[1] Moreover, for only 8 of the 22 innovations can this sluggish diffusion be rationalised by the classic vintage model argument that some existing old technology equipment was still earning positive rent; for the 14 other innovations there was no old technology embodied in existing equipment. At the very least, some major modifications to the profit maximising hypothesis are needed in order that it might analyse such a combination of facts. As Salter, himself, notes, entrepreneurs will often not act so as to maximise profits, if only because of uncertainty and 'the imperfect diffusion of technical knowledge' about new innovations.[2]

Consequently, the first two sections of this chapter survey the state of our knowledge in two relevant areas. Section 1 considers past empirical research on the evaluation methods actually used by entrepreneurs in deciding whether or not to adopt a new innovation. Section 2 summarises the few past studies concerned with the diffusion of information about new innovations. On the basis of this research, the most appropriate existing theory of the

1. cf. the government directives to nationalised industries to attempt to achieve an 8% rate of return on capital stock and the typical rate looked for by private industry of 15-16% on marginal, low risk investment ('The U.K.Economy,' ed. A. Prest and C. Coppock, Chapter 4 by J.R.Cable, p.168, 3rd ed.) See appendix one for information on typical pay-backs for the sample innovations.
2. W.Salter, (op.cit.), p.89. In fact Salter believes that even the modified hypothesis of profit maximising under uncertainty with imperfect knowledge is unlikely. He often refers to vigorous management, 'inertia' etc - concepts more in line with a managerial theory of the firm.

firm would appear to be the Behavioural theory as outlined by Cyert and March. Section 3 explores, briefly, the implications of this theory in the diffusion context; one conclusion being that each firm in a particular industry may be thought of as having an 'attitude' to each innovation. The influence on these attitudes of perhaps the most important industry-level characteristic - the degree of competition - is considered separately in section 4. Finally, section 5 examines the potential influence on the adoption decision of the innovation-supplying industry.

1. Evaluation methods - the evidence.

One way of examining the nature of the entrepreneurial adoption decision is to actually ask entrepreneurs about it. Such a direct approach to the problem does have well known drawbacks: non-response, evasive or misleading answers and interviewer bias in the way in which questions are posed. Invariably, the questionnaire or interview approach to decision-making produces the conclusion that firms do not behave in the relatively sophisticated manner suggested by conventional economic theory. Just as frequent, however, is the response of many economists that the rough, ad-hoc, decision-making usually reported in such research could be shown to be consistent with maximising behaviour were sufficient information available.[1]

It should come as no surprise, therefore, that there is little evidence from the surveys to be considered below that firms behave in an optimising way in deciding whether or not to adopt new innovations. The earliest study in the U.K. relates to 116 firms in the 1950's.[2] Carter and Williams asked these firms about the financial methods they used in deciding whether or not to invest in new innovations. Apparently, only 18 used some standard yardstick by which to judge whether adoption was worthwhile; for most of these, the

1. A classic example is the marginal cost - full cost pricing controversy stimulated by R. Hall and C. Hitch, "Price theory and business behaviour", Oxford Economic Papers 1939, vol. 2. See, for instance, F. Machlup, "Theories of the firm; Marginalist, Behavioural and Managerial," American Economic Review, March 1967.
2. C. Carter and R. Williams, 'Investment in innovation.' Oxford University Press, 1958 chapter 5.

standard was a pay-back period ranging from one to ten years; another thirty four claimed to make estimates of the likely costs and revenues, but said that no one pay back period was always used, or that they could not rely on these arithmetic calculations - 'commercial acumen' was also used. The remaining sixty four claimed to make no explicit calculations and acted rather on the basis of hunches or crude rules of thumb, such as whether their competitors had adopted.

Stoneman[1] reports some sample results on the evaluation methods used by firms when deciding whether or not to adopt computers. Of 360 firms sampled, only 71 responded,[2] of whom 14 used the pay back method, 18 used other methods (which Stoneman suggests were adaptations of the pay back) and 24 used the discounted cash flow (DCF) method. 25 used no method of evaluation and it is likely that an even larger proportion of non-respondents also used no formal method. As Stoneman points out, even payback criteria imply the principle of meeting targets rather than of obtaining a maximum return.

Past work on the evaluation methods used by British industry for more general investments suggest that both sets of above results are typical in the frequency of reported use of the pay back or no method at all.[3]

A survey(also reported by Stoneman) into firms' reasons for non-adoption of computers in the hotel and catering industry, yields the interesting finding that 62 of the 95 firms had not even explicitly considered the possibility, the other 33 claiming various other reasons of a rather general nature e.g.'too expensive.'

Whilst none of this evidence 'proves' non-optimising behaviour - given the nature of the research method, this would not be possible - it would require the construction of a very sophisticated maximizing model, with a

1. P.Stoneman, 'The choice of technique:The example of computerisation.' Warwick Economic Research Paper no. 46, 1974.

2. Some firms gave more than one answer.

3. For a summary of U.K. evidence, see P.Lund and D.Miner,'The investment behaviour of small firms.', part of Research report No. 11 to the Boulton Committee of Inquiry on small firms(Their summary is not confined only to small firms.)

large number of constraints, to rationalise such results. Alternatively, they can be interpreted as, by and large, indicating satisficing rather than maximizing behaviour. It will be conceptually useful, furthermore, to assume that each potential adopter has some yardstick measure, against which to judge its assessment of the returns from investing in a new process innovation. Even this much weaker assumption rests on the premise that, for some firms, this yardstick standard is implicit and, at the time of the above surveys, very high.

2. The diffusion of information about new innovations.

Of particular interest to a study such as this is the process by which information about new technologies spreads throughout the populations of potential consumers. Unfortunately, very little is known about information channels in industry and economists, as a group, have had little to offer theoretically or empirically in this area.

(a) First knowledge of the innovation's existence.

Of crucial importance is the typical length of time that elapses before all potential consumers are aware of the existence of the new innovation. Clearly, it is unlikely that all firms learn of the existence of a new innovation immediately it is made available by a machine maker. Indeed, an extreme statement that sometimes appears in the economic literature is that many firms are totally ignorant, even of the existence of the new innovation, for a number of years after its initial introduction. Carter and Williams[1] note that "the backward firm may not hear of an idea for several years after it is first made known," but it should be remembered that their interpretation of the word 'idea' is much broader than the major new innovations studied here.

1. C. Carter and B. Williams, "Industry and technical progress," Oxford University Press, 1957, pp. 178-9.

Nabseth[1] presents data for two innovations which, at face value, does suggest that some firms were totally ignorant of the innovation for a number of years after its first appearance. Firms were asked for the year in which they received their first information of the new innovations; for one innovation, the average time lag in six different countries, between first installation and all sample firms knowing of it, was five years; for the other innovation, the average lag for the same six countries was three years. On the other hand, the time lag between the first commercial application and 90% of sample firms knowing of the existence of the innovation was much shorter: for the first innovation, the average was less than three years, for the second, slightly more than one year. These are relatively short lags when viewed against the duration of the typical diffusion process: for none of the sample innovations in the present study had all firms adopted within ten years and, in most cases, it is unlikely that diffusion will have been completed even within twenty years.

Moreover, Nabseth's results are open to doubt on two counts. First, as he himself admits, the data used was "not precise; it is difficult (for firms) to recall exactly what happened fifteen years ago, even if - as in many cases - some of those (managers) who introduced the new process are still with the company."[2] Furthermore, for those firms unable to remember exactly when they first heard of the new innovation,[3] it is quite probable that their answers would be biassed towards a recent date. This might be especially so for some non-adopters. After all, what better excuse to give for not having adopted a new technique than that one was not aware of its existence until recently?

Secondly, it is not certain that all firms interpreted 'first information' in the same way. In some cases, the phrase may have been taken to mean

1. Concluding chapter of "New industrial processes", op.cit.(1974).
2. ibid., p. 300.
3. It would be surprising if they could remember something so intangible.

'first substantive information.' The distinction is not meaningless; 'first information' will often appear 'like manna from heaven,' substantive information, on the other hand, is more likely to result from a conscious search effort on behalf of the entrepreneur. Years later, he is much more likely to remember the latter than the former.

With one exception, little of relevance on this matter was uncovered in the technical literature on the sample innovations. The monopolist supplier of GA, ICI, apparently informed all potential adopters of the existence of its innovation once it was commercially available. It would be unlikely if this were the exception, especially as the population of potentially adopting firms for many innovations can be measured in tens rather than millions (as would be true for consumer durables.)

Total ignorance by a significant number of firms for any length of time, seems very unlikely, therefore. This is not to deny that, for many firms, knowledge may, initially, extend no further than awareness of the existence of the new innovation. Nabseth's data strongly implies that firms may have forgotten or ignored incoming information, presumably because they were unable to make much sense of it with respect to their own production methods. However, doubt as to the potential profitability of the innovations may not be due solely to the consumer's own lack of technical expertise; the innovation supplier may also be unsure, both because of general teething problems and due to the technical idiosyncracies of the consuming firm concerned.

(b) Sources of information.

It seems, then, that most firms learn of the existence of the innovation relatively quickly but that their initial knowledge may be patchy. Following on from this, how is this initial knowledge improved, and from what sources? The conventional view seems to be that most firms receive improved information from other members of their own industry. According to Ray, "Good reports of a new technique from entrepreneurs already using it may carry considerably more weight than reports in the press or publicity by suppliers."[1] Similarly, the only real reference made by Mansfield[2] to the quality of information available to the non-adopter, is that it improves as more of his fellows adopt (and presumably make their experiences available to him.)

Yet it is clear from advertising literature and from correspondence received from the innovation suppliers, that they divert substantial resources into marketing their innovations. In fact, Hakonson's data[3] confirms that the suppliers and trade journals were the most often quoted sources of information, given by potential adopters of the innovation he studied: over half of his respondents quoted them, as opposed to only infrequent mention of competitors as a source of information. Similarly, Rogers[4] claims that many (mainly sociological) studies support the hypothesis that impersonal information sources are most important at the 'awareness' stage.

Fortunately, some more interesting findings in this area are appearing in a continuing debate in certain market research journals. Webster,[5] for instance, interviewed the 'purchasing agents' of fifty New England manufacturing firms in a number of industries. Their general view of the best source of information about new products (which are, of course, the new processes or materials for the consuming firm), was the suppliers' own salesmen. They were considered by over half of the firms to be the most trusted source. Moreover, many respondents reported that they would divulge information about newly purchased products only with great reluctance, as they believed that the latter gave them a competitive edge. Further, the interviews provided little evidence for a bandwagon effect: only two of the fifty firms reported that certain other firms' purchases were important as an indicator of the worth of new innovations.

1. ibid. p. 9.
2. (1968), op.cit. pp. 135 - 6.
3. "New industrial processes.", op.cit., chapter 4.
4. op.cit., chapter 4.
5. F.Webster, "Informal communication in industrial markets,"Journal of Marketing Research, May 1970 pp. 186 - 9.

Similar findings are reported by Ozanne and Churchill in their study[1] of the adoption of an automatic machine tool in the U.S: a call from a salesman was often the instigator of a firm's first serious interest in the innovation. This was usually followed by an increased interest on receipt of a formal 'price quotation and tooling proposal.'

On the other hand, there is some evidence to suggest that inter-competitor information channels do exist. Martilla[2] reports that in the competitive U.S. greeting card industry, very little technical information is exchanged. But in the more localized, and thus less competitive, envelope industry, information was frequently shared among friendly firms. A similar picture emerges from Allen and Reilly's [3] study of the sources of technical information for Irish industry. Whilst interviewed firms claimed that they received little information from competing Irish or British firms, contacts with foreign firms were quite frequent - the latter, presumably, not being directly competitive with the Irish firms.

Perhaps the most surprising findings in this area, so far, are those reported in Czepiel's research[4] on the role of inter-firm contacts in the diffusion of continuous casting in the U.S. steel industry.[5] Firms were asked whether they had regular opinion/advice relationships with other firms in the industry. Apparently all but one of the thirty two firms had such contacts - on average, with two or three other firms in the industry. Rather surprisingly, direct informal interpersonal contacts among decision makers in different firms happened about once a week. More specifically,

1. U.Ozanne and G. Churchill, "Adoption research:information sources in the industrial purchasing decision," Proceedings Fall Conference, American Marketing Association, 1968, pp. 352 - 9.
2. J.Martilla, "Word of mouth communication inthe industrial adoption process," Journal of Marketing Research, May 1971, pp. 173 - 8.
3. T.Allen and V.Reilly, "Getting the word around", Report of a pilot study on technology transfer to Irish industry.M.I.T., Alfred Sloan School of Management. Working paper 650 - 73.
4. JCzepiel, "Word of mouth processes in the diffusion of a major technological innovation "Journal of Marketing Research, May 1974, pp. 172 - 9.
5. Not noted for a very marked degree of competition. See Adams and Dirlam's paper (op.cit.)

with respect to the particular decision to adopt the innovation, firms actively sought out information from other firms on top of these regular relationships. On average, before finally adopting, they approached 5.5 other firms, some of whom were contacted on many different occasions. In this study, however, no attempt was made to assess the importance of the suppliers' salesmen.

If Czepiel's findings are at all applicable to industry as a whole, it does seem unlikely, (a) that any firm will be totally unaware of the existence of any major innovation for very long and (b) that inter-firm contacts do not play any part in the diffusion of meaningful information. On the other hand, it seems plausible to suggest that both the frequency and information content of inter-firm contacts may be inversely related to the degree of competition prevalent in the industry concerned.[1]

Finally, there are two other reasons why technical information may be passed between competitors, either directly or indirectly, even when competition is real. First, some managers may derive a sort of 'kudos' from informing their peers of their experiences with a new, technically advanced, innovation. Second, non-adopters should be able to glean some information about the effects of the innovation, simply by observing any change in price or quality of the output of firms which have already adopted.

(c) Search as a conscious activity.

It is clear from much of the above research that information collection is not merely a passive activity on the part of the potential adopter. The search for information, by any one firm, may continue over a number of years, involving discussions with suppliers and, sometimes, competitors. Often, the sources of information may change as search becomes more rigorous. Both Martilla and Webster found that as search proceeded, an increasing number of opinions were sought (mainly from engineers, both inside and outside of the firms involved.)

1. Although Research Associations, as purveyors of information, may act as an important substitute for inter-firm contacts in some industries, e.g. the Shirley Institute in Textiles.

Interestingly, nearly always the individual(s) responsible for collecting and searching for information was an engineer, purchasing agent or production manager of some sort; in many cases they will not be responsible for making the financial decision to allocate the funds necessary for adoption to take place.

(d) Implications.

Whilst this (ongoing) research does provide some interesting insights, only a very partial picture has emerged as yet. Nevertheless, a number of tentative implications may be inferred.

At the firm level, ability to gain and process information may be partially determined by how cosmopolite are its managers. In this context, this may be measured by how receptive they are to trade journals, their propensity to attend industrial conferences and the level of their technical education. Hakonson,[1] following Carter and Williams,[2] has suggested, as proxies for such characteristics, the size of the firm's R and D department, the extent of its overseas interests and whether or not it belongs to a research association.

A popular view is that large firms may be more able than small firms to absorb and process information because they have larger engineering departments.[3] An alternative rationale suggested by the above discussion is that large firms are likely to receive more information from the innovation supplier because they present larger potential markets for the innovation.[4] Further, just because of their size, they may be in contact with more of their competitors. On the other side of the coin, information may be less efficiently diffused within the firm because of bureaucratic control loss.

1. S.Hakonson, op.cit.
2. Carter and Williams, (1957) op.cit., chapter 16.
3. E. Mansfield, (1968) op.cit., p. 156.
4. Empirical support is to be found in Nabseth's work (op.cit.)1973, in which large firms were found to receive, on average, 'first information' slightly earlier than small firms.

At the industry-level, as has been suggested, the quality and extent of information exchange may be inversely related to the degree of competition. But it may also be inversely related to the number of firms in the industry. As Williamson points out[1], it becomes increasingly difficult to keep up comprehensive contacts with the rest of one's industry, the more firms there are in that industry. As N (the number of firms) increases, the volume of 'transmissions' necessary to keep each firm in touch with every other increases roughly in proportion to N^2. Similarly, the efficiency of information transmission by the innovation producer must decline. For instance, a market of 10 consumers, each with a potential purchase of 10 units of the innovation, can be more easily covered by a salesman than a market of 100 consumers, each with a potential purchase of 1 unit. Thus a large N might act as an important drag on the diffusion of information if the producer has a fixed supply of salesmen (typically, trained scientists or technologists.)

Certain characteristics of the innovations, themselves, may influence information flows. It is possible that a very profitable innovation will tend to be 'talked about' more often than others, thus 'word will get around' much more quickly and non-adopters may be encouraged to allocate more resources to search. Conversely, high profitability may lead to adopters keeping their experiences to themselves.

The technical sophistication of the innovation might be expected to influence the extent and returns to search. Using the classification developed in section 7 of the previous chapter, group B innovations may require extensive search before information is complete. Moreover, given the continuing post-invention improvements to their specification, search will continue to yield new information for a relatively long period. For class A innovations, however, much less search will be required and it will be subject to diminishing returns at a much earlier stage.

Finally, the innovation suppliers will also determine the quality of information diffusion. This possibility will be considered shortly, within an overall assessment of the role of the supplying industry.

1. O.E.Williamsom, "A dynamic theory of inter-firm behaviour," Quarterly Journal of Economics May 1965.

3. An underlying theory of the firm.

To summarise, briefly, in sections 1 and 2, certain features of the adoption process have been suggested. In deciding whether or not to adopt new process innovations, firms use various rules of thumb based on yardsticks, requiring little financial sophistication. Generally, perfect information about innovations does not occur. Whilst most firms 'hear of' the innovations relatively quickly, the extent of their knowledge is initially limited, only improving with time due to 'active' and/or 'passive' search. Some improvements in knowledge are virtually costless, resulting from advertising,(in some form), from the suppliers and social intercourse with competitors who have already adopted. However other improvements, possibly of a more substantial nature, do incur active search and thus certain costs; repeated discussions with suppliers and previous adopters often appear necessary. These discussions and technical evaluations may extend over a number of years and certainly exceed what might be called the normal level of social intercourse. Finally, it is probable that the extent of search behaviour and information receptiveness will vary depending on certain characteristics of the firms, industries and innovations involved.

In order to use this description as the basis of a determinate theory of decision making, certain aspects must be developed; particularly, the motivation behind search and the determinants of the yardsticks used in evaluation.

(a) The behavioural approach.

One plausible approach to these problems is to assume that firms may best be described on behavioural lines. Certainly many of the findings of the previous two sections are in accordance with the 'Behavioural Theory of the Firm' as outlined by Cyert and March.[1] The aspects of this theory most

1. R.M.Cyert and J.G.March, 'A Behavioural Theory of the Firm'. New Jersey, Prentice Hall, 1963. This is not to say that other rationalisations are not also possible. It could be argued, for instance, that none of these findings is inconsistent with constrained profit maximising. If there are significant costs in evaluation, it is conceptually feasible that the use of rules of thumb in decision making may be optimal behaviour for profit maximisers. Similarly, if there are significant search costs, a state of non-perfect

continued ...

pertinent in this context can be summarised fairly briefly. The firm is viewed as a coalition of different factions, each with potentially conflicting interests (e.g. shareholders, financial managers, technical managers, workforce etc.) The everyday running of the firm is based on the pursuance of certain goals (for instance, sales, profits and market shares goals.) Search is activated when these goals are threatened or unfulfilled. Thus it is problem-oriented; moreover, initially, it will be simple-minded and localised; managers may only look outside their own firms if initial search is unproductive. Decision making is based on satisficing rules: a course of action will be pursued if it is expected to satisfy certain minimum criteria.

In the present context, then, the new innovation is a potential solution to the non-attainment of goals (but also, perhaps, the reason for their non-attainment.) Its probable success as a solution will be judged by comparing the expected outcome from adoption against some minimum yardstick performance which must be satisfied (e.g. a minimum required payback.) If the innovation is initially rejected as a solution, a number of possibilities arise: (a) it may be dismissed as an alternative until some later date when improved incoming information re-directs attention back to it, (b) search for more information may be activated, perhaps implying approaches to other firms and the suppliers, (c) the standards which it is asked to satisfy are relaxed. Thus, a firm strives for a solution 'either by discovering an alternative that satisfies the goals or by revising the goals.'[1]

(footnote 1. continued from previous page...)
information about a new innovation may be optimal. Unfortunately, the predictive power of the neo-classical theory of the firm is inversely related to the number of constraints added. Given sufficient constraints, the neo-classical theory will surely eventually resemble the behavioural theory quite closely, but only at the cost of losing its traditional simplicity.

1. ibid. p.121.Strictly speaking, goals should not be confused with rules of thumb.However, in certain cases the distinction is very fine. For instance, one goal of the firm may be the attainment of a rate of return of x% on all capital; at the same time, the decision to invest may depend on the potential investment being required to satisfy a minimum return of y%. Thus, x% is the goal and y% the yardstick used in the rule of thumb. At any event, however, a relaxation in the goal would surely lead to a similar relaxation in the yardstick.

(b) The yardstick used in investment appraisal.

Generally, but not always, or even uniquely, the yardstick used in this area will be some critical rate of return or pay-back,[1] which must be expected to obtain for adoption to be acceptable. There are good reasons to suppose that this critical rate will vary across firms within the same industry.[2] Loosely, it may be conceptualised as representing the firm's attitude to the new innovation. As such, it may well be determined by implicit bargaining between the various factions of the coalition, and will thus reflect the importance to individuals of various goals and also the bargaining strength of those individuals. Thus, a firm in a research intensive industry with scientists or engineers on the board of directors might set a lower critical rate than might a family business in an old craft-industry.[3]

One particularly interesting possibility is that the critical rate may be correlated with the size of firm. The arguments for and against large firms having a greater tendency towards 'progressive' attitudes are, by now, well known.[4] Large firms might be expected to employ more scientists and engineers and have easier access to finance; further, should the innovation 'fail', the proportionate consequences will be less serious, the larger the firm's overall operations. On the other hand, large firms may be more bureaucratic and thus, perhaps, more conservative. In behavioural terms, large firms might be more likely to use lower critical rates because they usually have stronger technical lobbies within their management; their

1. In some cases, particularly family businesses, this may not be explicit, even then the concept of satisfying some minimum requirement may still be in evidence, however.
2. But not in the limiting case of perfect competition in which the behaviouralist firm is forced, by the 'hidden hand', into a profit-maximising posture.
3. In extreme cases, new innovations may actually be resented. In his study of numerical control (NC) in the U.S. tool and die industries Mansfield reports that some firms were on record as claiming that NC was a threat to the craft nature of their industry. Mansfield et al, op.cit., 1971.
4. Mansfield (1968) op.cit., provides a useful list of references on this point, p. 156 footnote 2. A comprehensive survey is to be found in F.M.Scherer, 'Industrial Market Structure and Economic performance'. Rand, McNally, Chicago, 1971, chapter 15.

financial lobbies may be less antagonistic since the cost of finance may be lower, and the innovation poses less of a potential threat to the other goals. Alternatively, the larger the size of the firm, the more likely is decision making to involve more conflicting interests which may lead to conservative, compromise yardsticks.

Another distinct possibility, which has received little attention in this particular debate, is that the ownership-control split may influence decision making. Whilst adoption of a new innovation may be in the interests of one faction in the coalition - the shareholders - it may be less attractive to other factions - the managers (particularly production managers.) More often than not, the quality and effort required of management by new processes is higher than for the older technologies. Problems of loss of flexibility in product mix and the added significance of high capacity working have been mentioned already. In addition, raw materials often have to conform to tighter specifications, expensive quality control is sometimes necessary, wage payments tend to become more rigid and extra and highly skilled staff often need to be employed.[1] Generally, then, adoption may lead to extra profits for the owners (shareholders) but significant re-organisation for managers. Where management are not also the owners of the firm, the relative position of the owners in the coalition will presumably be weaker. This may lead to greater relative importance being attached to the disruptive effects of the new process. In turn, the critical rate, used to assess adoption, may well be higher than would be the case if there was no owner-control split. To the extent that larger firms tend not to be owner-managed, therefore, there is a possibility that they will use higher critical rate yardsticks.

A counter argument, however, is that smaller owner-run business may be loathe to adopt, if to do so would require the introduction of new skilled

1. See F.Eels, A.Hazlewood, K.Knowles and C.Winston, 'Innovation and Automation,' Bulletin of the Oxford University Institute of Technology Statistics, Vol. 21 (1959) pp. 131 - 208.

management who might dilute the owners' power base.

Needless to say, many other hypotheses have been suggested in the literature as to the determinants of firm-attitudes: education and age of management, the degree of internal financing, profit trends, growth rates etc.[1]

The technical complexity of the innovation will obviously influence firms' critical rates. There is always a risk with any new innovation that it will not function properly technically and that, even if it does, the expected cost improvements do not live up to expectations. Both risks (presumably more real the more complex is the underlying technology) will undoubtedly be reflected in a risk premium to be added to the normal critical rate. A less obvious point concerns innovations undergoing rapid post-innovation improvements. Under these conditions, the prospective buyer may have a problem of timing - although the innovation may already have been seen as sufficiently profitable at time t, by waiting until time t + 1, an even more profitable vintage may be purchased.[2] In such a case (most likely where the technology involved is quite complex, thus resulting in considerable learning by doing by the manufacturers) the decision maker may require an extremely high present expected return to dissuade him from holding back.

A number of industry-level characteristics may also be of some importance. For instance, the less competitive is an industry, the more firms will be able to pursue goals other than profits, and the more scope there is for the use of conservative[3] rules of thumb; furthermore, there is likely to be greater diversity of the critical rates used, if only because of the many alternative goals which may be pursued.

1. See, for instance, the discussion of section 3 of the 2nd chapter; also Carter and Williams, op.cit. (1957), chapter 16, particularly. These possibilities will not be pursued here because of the lack of widespread data on such variables. They have hardly ever proved significant in past research-probably due to measurement difficulties.
2. However, the simple minded search of the typical behavioural firm implies a lower probability of such waiting than would the conventional theory of the firm.
3. Although some goals (e.g. growth) might lead to the use of a less stringent critical rate.

Similarly, demand conditions may also be influential. Greater suspicion of risky techniques might be expected where demand is static or falling and where fluctuations in demand are, typically, large.

(c) Dynamic aspects of the adoption decision.

Within this behavioural schema, there are two factors which increase the probability of adoption with the passage of time. Search improves the quality of information available about the innovation and reductions in the critical yardstick rate increase the chance of adoption, for a given level of information. Conceptually, improved information resulting from search increases the chance of adoption in two ways: it may either improve the firm's view of the returns to be gained from adoption and/or increase the confidence with which that view is held (in turn reducing the risk premium.) The critical rate may be reduced either because goals are unfulfilled or, as just suggested, because the risk attached to adoption declines.

A useful distinction which may be made is between the exogeneous and endogeneous[1] reasons for these changes over time. Exogeneous influences, such as labour shortage, a slump in demand, pressure on existing capacity, may threaten certain goals within the firm. In response, the firm will search for solutions and, if these are not forthcoming, given existing goals or rules of thumb, then the latter may be revised downwards. Of course, adoption of the new innovation may be only one of a number of alternative solutions. Intuitively, the more reputedly profitable the innovation, the greater the chance that it will be the solution to which most search is allocated.

In addition, other influences may be directly attributable to the innovation

1. Basically, endogeneous influences might be attributable to the innovation itself. Exogeneous influences are more general, and would occur even if the innovation did not exist. The distinction is a little forced but it will prove to be convenient for the exposition in chapter 5.

itself. The most theoretically interesting endogeneous influence is the mounting competitive pressure on non-adopters arising from competitors having adopted.[1] Again, the response of the non-adopter should be to increase search and perhaps reduce the critical rate used to assess the investment decision. Secondly the risk attached to non-adoption will decrease for many firms, just with the passage of time: they will observe other firms using the innovation with apparent success (especially if adoption is seen to accompany a reduction in price or an improvement in product quality.)

Thirdly, the very existence of the innovation may suggest a new goal for the technical members of the management coalition: certain managers may derive utility from being 'technically progressive' and, thus, adoption of a new innovation may be seen as a manifestation of such progressiveness. Pursuit of this new goal may produce search, even if no existing goals are threatened.

Thus, both exogeneous and endogeneous influences will improve the chance of adoption because of increased search. However, whilst endogeneous influences will tend to persuade firms to reduce the critical yardstick rate, exogeneous influences will sometimes have the opposite effect. When the state of the world is favourable, goals will be achieved and as Cyert and March suggest, this may lead to an upward revision of goals and yardsticks.

Briefly, there are two major implications. For many non-adopters, the probability of a positive adoption decision increases with the passage of time, due to endogeneous factors. Over and above this, exogeneous factors may increase or decrease this probability, depending on the state of the world facing the industry. Moreover, it seems probable that the exogeneous influences will be cyclical; most of the examples quoted above (labour shortage, for instance) do have pronounced cycles. Whether or not the net effect of the business cycle is positive is not certain. Some goals are more attainable under boom conditions (e.g. high profits and growth), whilst others may be

1. Accordingly the following section is allocated to a separate consideration of this influence.

easier to satisfy under slump conditions (e.g. avoidance of labour shortage and capacity shortages.)

4. The role of industrial structure.

So far the level of competition has appeared in two contexts. First, as described in the last section, the attitudes of non-adopters will change as diffusion proceeds because they come under increasing pressure. As more and more of their competitors adopt, their increasingly deteriorating competitive position leads to the non-attainment of certain goals. This, in turn, produces a re-assessment of goals and/or increased search for solutions. Both factors may lead to more favourable attitudes towards the new innovation. Second, in section 2, some evidence was cited which suggests that the returns to information search might be inversely related to the level of competition in the industry. A third possibility, not considered so far, is that the actual returns from adoption may also be partly determined by the competitive structure of the industry.

This section will be devoted primarily to the first and third of these possibilities. Specifically, are the competitive pressures on non-adopters and the returns from adoption greater, the more competitive the industry in which the innovation is diffusing?

(a) Perfect competition versus monopoly.

Salter shows that there is 'no reason for a greater delay in the introduction of new techniques in monopolistic industry compared to competitive industry,'[1] so long as firms are profit maximizers. This conclusion flows from the equalisation, in both cases, of total costs of marginal new capacity with the operating costs of marginal existing capacity. "However, the important difference is that the monopolist is under no external pressure (that is, other than his own self-interest) to scrap obsolete equipment; while the producer in a competitive industry is forced to do so by the price changes resulting from the actions of his competitors."[2] The appearance of a new technique

1. Salter, op.cit, p. 93.
2. ibid., p. 93.

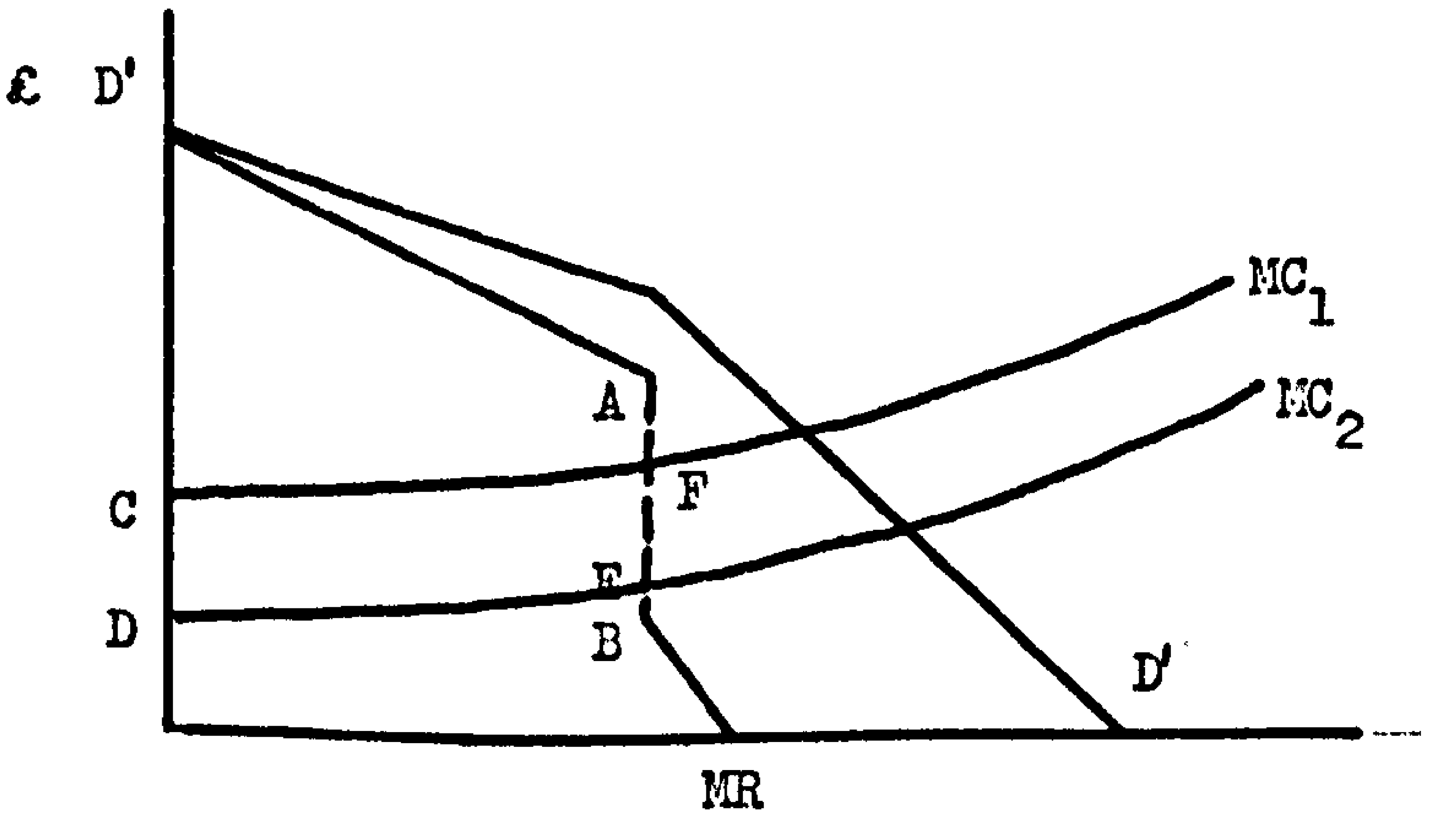
with lower costs, then, presents the same incentive – extra profits – but differing pressures. This analysis was based upon the classic vintage assumptions, including indivisible plant, and whilst it becomes less clear cut when allowing for piecemeal changes in plant[1] the main conclusion still stands.

ii) oligopoly and intermediate market structures.

On an intuitive basis, there is some disagreement as to whether oligopoly encourages rapid adoption of new techniques or not; Adams and Dirlam[2] suspect that oligopolists may tend to refrain from pioneering for fear of upsetting the status quo, but say nothing about the reactions of other firms, once a new technique has been pioneered. Salter, on the other hand, suggests that the state of oligopoly places a premium on being one jump ahead with respect to new techniques, 'so as to ensure the expansion of output, which these allow, is achieved by his new capacity rather than that of his competitors (thus avoiding direct aggression.)'[3]

A useful theoretical construction which distinguishes the pressures and incentives resulting from the advent of a new process innovation is the kinked demand curve.[4]

Figure 4.4.1. The kinked demand curve: Incentives to adopt.



1. which means that it is impossible to allocate rent to any one part of the now divisible plant. See the earlier discussion of this point in chapter 3.
2. op.cit.
3. op.cit., p.93.
4. See G.Stigler, 'The kinky oligopoly demand curve and rigid prices,' Journal of Political Economy, October 1947, for a general criticism of the kinked demand curve.

In the diagram, as usual, D'D' is the demand curve facing the firm and D'ABMR its marginal revenue curve. If adoption of the new process would reduce the marginal cost curve from MC_1 to MC_2, then there would be no incentive to reduce price or increase output; if price were reduced by the extent of the cost savings, all competitors would reduce their prices accordingly and only a small increase in the firm's demand would take place. The reaction of the adopter would be, therefore, to leave product price unchanged. This does not mean that he has no incentive to adopt - quite obviously he does, his profit increasing by an amount equal to the area of CDEF. But, his adoption does not affect his competitors at all. They lose none of their market share and are under no pressure to adopt, no matter how many of their competitors have adopted. There are, however, two exceptional cases. If the cost savings are large enough to reduce MC_2 to such a level that it cuts the MR curve below the point B, then there is an incentive for a profit maximiser to reduce his price accordingly, thus putting pressure on non-adopters. Secondly, if the new innovation improves the quality of the product, then non-adopters would be under pressure to follow suit, as they would be now offering an inferior quality, if not an inferior price.

(iii) Average cost pricing and product differentation.

Given the theory of the firm followed in section 4, it would be quite inconsistent to base the analysis in this section entirely on models incorporating profit maximisation. Consequently, a brief analysis of a non-maximising average-cost pricing model is provided, in the hope that the above conclusions might be somewhat generalised. It must be acknowledged that the substitution of average cost for marginal cost pricing is not sufficient reason for claiming a behavioural analysis. (Indeed, the behavioural approach can rarely be translated easily into the formalism of mathematics.) Nevertheless it is instructive to examine the implications of relaxing the optimising assumption. Moreover, price setting at some mark-up over average cost is broadly consistent with the behavioural view of decision-making based on crude non-optimising rules of thumb.

Suppose that in an industry of N firms the demand curve facing firm i is given by:

$$q_i = A\, p_i^{-\eta_i} \prod_{j \neq i}^{N} p_j^{\eta_{ij}} \qquad (4.4.1.)$$

where q_i is the demand for firm i's product, p_i its price and p_j the price of its j th competitor, η_i is its own price elasticity, η_{ij} the cross price elasticity of its demand with respect to j's price, and A reflects non-price determinants of demand such as incomes and tastes.

Thus the industry is characterised by product differentiation, each firm facing a downward sloping demand curve with demand also being sensitive to other firms' prices.

Firm i sets its price equal to some unknown multiple $(1 + M_i)$ of average costs (AC_i) where M_i is a constant mark-up[1]:

$$p_i = AC_i\,(1 + M_i) \qquad (4.4.2.)$$

The profit function is given by:

$$\pi_i = q_i\,(p_i - AC_i) = q_i M_i AC_i \qquad (4.4.3.)$$

The change in profits from a fall in average costs due to adoption of a new innovation is $d\pi_i$, where

$$d\pi_i = dAC_i \left\{ M_i q_i + M_i AC_i \frac{\partial q_i}{\partial p_i} \cdot \frac{\partial p_i}{\partial AC_i} \right\}$$

$$= dAC_i \left\{ M_i q_i - \eta_i q_i M_i \right\} = dAC_i M_i q_i\,(1 - \eta_i) \qquad (4.4.4.)$$

Thus, since $dAC_i < 0$, so long as $\eta_i > 1$, it will always be profitable to adopt.

The returns from adoption will always be greater the less monopoly power firm i has (or the bigger is η_i), the larger its existing output[2] and the greater the

1. There can never be a precise specification of M_i in average cost models: it might be related, however, to the barriers to entry into the industry, and the relative bargaining positions of the various factions of the management coalition.
2. Total increase in profits may be higher the larger is the firm, but, unless there are economies of scale in the price of the innovation (see previous chapter), the immediate rate of return will be insensitive to scale.

reduction in average costs effected by the innovation. On the other hand, returns are proportional to the size of the mark-up (M_i) which, as noted in the footnote, may be <u>inversely</u> related to the level of competition. Without specifying the determinants of M_i, therefore, it is not certain whether <u>incentives</u> to adopt are higher or lower in more competitive industries.

In order to use this model to analyse the <u>pressures</u> on non-adopters to adopt, a number of non-crucial assumptions may be made which considerably simplify the exposition.

Let all non-adopters have the same average costs (AC_o) and all adopters the same, lower, average costs (AC_N).

Further, if all firms use the same mark-up $\bar{M}$, have the same own price elasticity $\bar{\eta}$ and cross price elasticites $\bar{\eta}'$ and all demand functions are homogeneous of degree zero; then:

$$\begin{aligned} M_i &= \bar{M} \quad \text{for all i.} \\ \eta_i &= \bar{\eta} \quad \text{for all i.} \\ \eta_{ij} &= \bar{\eta}' \quad \text{for all i and j, } j \neq i \\ (N-1)\,\bar{\eta}' &= \bar{\eta} \end{aligned} \tag{4.4.5.}$$

Assuming i is a non-adopter at time t, his demand function may be written as:

$$q_{it} = Ap_{it}^{-\bar{\eta}} \prod_{j \neq i}^{N} P_{jt}^{\bar{\eta}'} \tag{4.4.6.}$$

If, at time t, there are m firms having adopted the innovation, using (4.4.2.), (4.4.6.) may be written as:

$$q_{it} = A\left[AC_o(1+\bar{M})\right]^{-\bar{\eta}}\left[AC_o\ (1+\bar{M})\right]^{+(N-m_t-1)\,\bar{\eta}'}\left[AC_N(1+\bar{M})\right]^{m_t\bar{\eta}'}$$

or $$\begin{aligned} q_{it} &= A\left[AC_o(1+\bar{M})\right]^{-\bar{\eta}+(N-m_t-1)\bar{\eta}'}\left[AC_N\ (1+\bar{M})\right]^{m_t\bar{\eta}'} \\ &= A\left[\frac{AC_N}{AC_o}\right]^{m_t\bar{\eta}'} \end{aligned} \tag{4.4.7.}$$

Clearly, since $AC_N < AC_o$ and m_t rises monotonically with t, q_{it} must decline over time. Differentiating with respect to time,

$$\frac{dq_{it}}{dt}\frac{1}{q_{it}} = \frac{dm_t}{dt}\,\bar{\eta}'\log\left[\frac{AC_N}{AC_o}\right] < 0 \text{ as } \frac{AC_N}{AC_o} < 1 \tag{4.4.8.}$$

Thus, there is a contraction in sales for the non-adopter related to the change in diffusion and the typical cross-price elasticity in the industry. In this case, then, <u>pressures will be greater the more competitive is the industry</u>. If there is also a continual decline in AC_N due to post-invention improvements in the new innovation,[1] the dynamic pressures will be re-inforced.

One of the limitations of this model is the indeterminacy of M_i. This problem is removed however, if one is prepared to substitute for (4.4.2.) the profit-maximising assumption:

$$p_i = MC_i \left(\frac{\eta_i}{\eta_i - 1}\right) \qquad (4.4.2A.)$$

(derived from equating MC_i and MR_i).

Whence $$\pi_i = q_i \left(MC_i \frac{\eta_i}{\eta_i - 1} - AC_i\right) \qquad (4.4.3A.)$$

and $$d\pi_i = dAC_i\ q_i \left[\frac{dMC_i}{dAC_i}\ \eta_i \left(\frac{AC_i - MC_i}{MC_i}\right) - 1\right]$$

Assuming for simplicity that $\frac{dMC_i}{MC_i} = \frac{dAC_i}{AC_i}$

$$d\pi_i = dAC_i .\ q_i \left[\eta_i \left(\frac{AC_i - MC_i}{AC_i}\right) - 1\right] \qquad (4.4.4A.)$$

Now, the extent to which average cost exceeds marginal costs at the level of initial output is a measure of the barriers to entry into the industry. In the limiting case of free entry, price is driven down to equality with average cost,[2] and $MC_i\left(\frac{\eta_i}{\eta_i - 1}\right) = AC_i$

In this case, $d\pi_i = dAC_i .\ q_i .\ 0 = 0.$

Generally, however, barriers will exist which will permit some degree of super-normal profits to be earned by existing members of the industry.

If $p_i - AC_i = f(B_i)$

where B_i reflects the height of entry barriers, and $f(B_i) \geq 0$, $\frac{\partial f}{\partial B_i} > 0$

1. See previous chapter, section 7.
2. The traditional theory of monopolistic competition, which allows for product differentiation <u>and</u> free entry, see E.H. Chamberlin, 'The theory of monopolistic competition,' (7th edition), Cambridge (Mass.), Harvard University Press (1956).

then $\frac{\eta_i}{\eta_i - 1} MC_i - AC_i = f(B_i)$

substituting into (4.4.A),

$$d\pi_i = \frac{dAC_i}{AC_i} q_i \left[\eta_i AC_i - (\eta_i - 1)f(B_i) - (\eta_i - 1)AC_i - AC_i\right]$$

$$= \frac{dAC_i}{AC_i} q_i (1 - \eta_i) f(B_i) \qquad (4.4.4B)$$

As $(1 - \eta_i)$ must be negative for positive profits ever to occur and $\frac{dAC_i}{AC_i} < 0$, adoption will always be profitable. Once again, incentives will be higher the greater the proportionate saving effected by the innovation, the larger the firm's size and the more elastic its demand ; however, this time high entry barriers will also produce higher incentives. Thus a more precise prediction is possible if firms are profit maximisers.

Again, making assumptions (4.4.5), the demand facing the non-adopter at time t is given by

$$q_{it} = A\left[\frac{MC_N}{MC_o}\right]^{m_t \bar{\eta}'} \qquad (4.4.7A)$$

Thus, the conclusion that pressure is more intense the higher is $\bar{\eta}'$ (and thus the more competitive is the industry) is maintained.

In conclusion then, the conventional proposition may be accepted on the basis of the above analysis. Clearly, competitive pressures on non-adopters should be more intense the more competitive is the industry concerned. In the monopoly and standard kinked demand curve cases, the only penalty incurred by non-adopters is a profit level below that which could be achieved. In the product differentiation models, (whether assuming average cost or marginal cost pricing), demand for the non-adopter's product declines at a rate dictated by its price elasticity of demand – the more differentiated is the firm's product, the slower will be the decline in demand. In the extreme case of perfect competition, non-adoption evokes the ultimate sanction – loss making.

In terms of the behavioural analysis, these pressures will, to a greater or lesser extent, gradually increase the probability of non-attainment of goals: rate of return on capital, overall profits, market share and sales may all be reduced. Thus the extent of search and/or the willingness of non-adopters to reduce their targets and yardstick rules of thumb will be partly determined by the degree of competition in the industry concerned.

Turning to the incentives to adopt (as reflected by the returns to be gained from adoption), the position is far less clear. There is little in the conventional analysis of monopoly, perfect competition and the kinked demand curve to suggest that incentives will be smaller in less competitive industries. On the other hand, the analysis of product differentiation yields a confusing picture: incentives should be higher the more elastic is the potential adopter's demand curve, and the greater the barriers to entry into his industry. However, high own-price elasticities are usually associated with competitive industries whilst high entry barriers are associated with more oligopolistic industries.

Finally, a reservation should be added. For many of the sample innovations, adoption often produces an improvement in the quality of the end-product. Quite significant product quality improvements occur when VM, VD, NC, SPC and CT are used, and lesser improvements often occur after adoption of SP, F, SF, WSB, PCBC, SL, ASB and BOP.[1]

In such cases, the incentives and pressures will occur not only because of the cost-advantages of the new innovations. Adoption may result in increased demand because of improved quality,[2] as well as reduced price. Similarly, non-adopters will lose part of their market share partly because they are offering an inferior quality product.

Thus, even in the most price-rigid of market structures (e.g. oligopoly facing a kinked demand curve, price leadership or collusion), quality differentials may produce powerful pressures on non-adopters.

1. Thus, strictly speaking, product, as well as process innovation is involved.
2. For instance, E.Mansfield et al (1973) op.cit., reports that the competitive pressures on precision machining firms to adopt numerical control were much greater than for those non-precision machining in the U.S. tool and die industry.

5. The role of the innovation supplier in the adoption decision.

As indicated in the previous chapter, all 22 of the sample innovations are sold or installed by a capital goods industry which has usually been responsible for the original invention and development work. Yet, almost without exception, past research into diffusion has totally ignored this possibility, both theoretically and empirically.[1] Whilst most of the interesting aspects of diffusion do concern the consumers' decision making behaviour, any analysis which totally ignores the role of the innovation producer is, at best, only partial.

Three aspects of the supplier's behaviour will affect his consumers' adoption decisions: his research expenditures and pricing policy will obviously determine the profitability and costs of the innovation and his advertising expenditures will influence the quality of information available to the consumer. For these reasons, it is unacceptable to define diffusion as only a demand phenomenon, assuming that supply is a passive onlooker.

Having said this, data collection in this area has proved almost impossible in this study. The innovation manufacturers are usually prepared to supply advertising literature, and, sometimes, quite useful qualitative data about technological developments, but, not surprisingly, they are rather less willing, or able, to provide more quantitative information.

Nevertheless a brief summary of the characteristics of the supplying industries does provide some important findings.

(a) The suppliers of the sample innovations.

Contrary to expectation, the sample innovations are not generally supplied by monopolists.[2] Only four (all supplementary and cheap) are produced by one firm: GA, SF, EH and ASB; for ASB and EH monopoly power was allocated under licence to the firm in question by the Research Association responsible

1. But see Stoneman (1974), op.cit., for a lengthy discussion of the role of the manufacturers in the spreadof computers in the U.K.
2. Although, initially, of course, most genuinely new innovations will be supplied by only one firm.

for the innovations, ICI's monopoly of GA appears to be the result of patents and superior technical 'know-how.' The vast majority of the sample innovations are supplied by between two and eight firms; only two (NC and ATL) are produced by more than eight firms. Oligopoly may therefore be defined as the typical case. (Invariably these 'industries' constitute parts of various 'minimum list headings' in the Mechanical Engineering industry.)

A general feature of these industries is their international nature: for fifteen of the twenty two innovations, the supplying industry might be said to be an international oligopoly, in the sense that the firms selling the innovations to the consumers are internationally functioning, or are subsidiaries of foreign firms. This is to be expected, of course, given the international nature of many of the inventions. Consequently, it is likely that the suppliers will have ready access to improvements in the technology emanating from overseas.

The other notable aspect of most of these industries is their close historical links with the consuming industries. For nearly all of the innovations, producers are traditional suppliers, having supplied many previous innovations. In at least nine cases, the consuming industries constitute the main, if not the sole, market for the suppliers. For instance, the producers of tufted carpet machines supply a wide range of other products predominantly to the carpet industries. For only three innovations (GA, CT and PCBC) could the innovation producer be said to be breaking into a new market. Without more data the implications of this are uncertain. It is probable, nevertheless, that information should diffuse relatively quickly with personal contact having already been established in the past.[1] It seems unlikely that potential consumers will be unaware, for very long, of the existence of new innovations. Certainly most of the advertising literature

1. This may reduce the need for formal advertising. Interestingly, formal advertising expenditure is usually found to be much lower in the capital goods industries than in the consumer industries. See Pratten, op. cit., chapter 31.

tends to emphasise the merits of the particular brand of the innovation rather than the generic advantages.

(b) A stylised description of the supplier's role.

From this brief survey and the previous discussions of learning by doing and information diffusion,[1] a stylised description of the innovation supplier and his potential role becomes possible. Typically, the innovation is supplied by an oligopoly (but sometimes a monopoly) which is already in close contact with its market. For technological reasons the physical inputs required to produce the innovation will decline over time,[2] whilst the quality of the innovation will improve over time. Furthermore, suppliers are frequent and influential purveyors of information about their innovations.

Of central importance, are the implications for the pricing of the typical new innovation, but, in addition, the previous discussions of post-invention improvements to the innovations and information diffusion may now be developed a little further.

(c) A model of pricing.

There exists, in the literature, no theoretical model of oligopoly pricing of new industrial products and very little of relevance is available on monopoly pricing. Of somewhat peripheral interest, is the interchange of articles and comments between Arrow, Demsetz and Yamey;[3] their controversy concerned the incentives for a monopolistic inventor to supply a new innovation to a monopolistic consumer, as opposed to a perfectly competitive consuming industry. A consensus view did not emerge, but the most convincing arguments do appear to be on the side of the competitive industry offering bigger incentives.

1. Section 7 of the previous chapter and section 2 of this chapter respectively.

2. Because of the international nature of many firms, the learning curve is best defined, for the U.K., with respect to time rather than cumulative output, as the latter may well depend on overseas production.

3. See C.K. Rowley, 'Antitrust and Economic Efficiency' MacMillan, London 1973, for a fairly comprehensive survey of this controversy.

A recent model due to Glaister[1] describes the optimal time path of price and advertising expenditure, for a monopolistic seller with a new product. Demand is assumed to follow a logistic growth curve[2] with its slope parameter (β in terms of equation 2.1.3.) being a function of product price and advertising expenditures. Under these circumstances, Glaister shows that a long-term profit maximising monopolist[3] will continuously revise price such that it, too, follows a logistic-type curve over time. This policy involves selling a number of units at below their production costs, increasing price only slowly and only attaining the conventional monopolist's price after a considerable time. A similar sort of prediction emerges for the optimal advertising strategy.

Needless to say, his analysis would become virtually impossible mathematically if certain assumptions are relaxed. For instance, the twin assumption of constant costs and elasticity of demand are crucial. Yet, as suggested in the previous chapter, costs are unlikely to be constant. Even more important, however, like all optimising models of this sort[4] it is <u>pre</u>scriptive rather than <u>de</u>scriptive: it is doubtful whether firms are capable of operating so as to maximise their discounted net worth. Indeed, Glaister himself admits to firms not operating according to his model's predictions.[5]

Clearly for any dynamic model of pricing to yield testable predictions, a fair degree of abstraction is necessary. In this case, certain aspects

1. S.Glaister, 'Advertising policy and returns to scale in markets where information is passed between individuals,' Economica, May 1974, pp. 139-156.
2. Whilst this is plausible for new non-durable products for which the model is developed, it is less likely for durables or process innovations in which replacement buying is far less important. In these cases a bell shaped curve is much more likely: whilst the number of owners of the new process may grow according to an S-shape, the number of new purchases will not.
3. One who maximises his net worth into the future, subject to his demand curve.
4. For surveys of such models, see Glaister, ibid, and A.Jacquemain and J. Thisse, 'Strategy of the firm and market structure : an application of optimal control theory,' in 'Market structure and corporate behaviour,' ed. K. Cowling, Gray-Mills London 1972.
5. op.cit., p. 154.

of the problem merit particular attention: for instance, it is not really satisfactory to assume constant cösts over time. In the model to be presented below certain simplifying assumptions are made on the less crucial aspects in order that the implications of the stylised description offered above may be examined.

The typical supplying industry, being oligopolistic and having close traditional links with its consumers, is likely to be characterised by product differentiation. Such a situation can be most appropriately analysed assuming either average cost pricing or marginal cost pricing subject to a downward sloping demand curve.

Let all suppliers have a production function with only one input (labour)[1] and constant returns to scale at every point in time:

$$MC_{it} = AC_{it} = w_t L_{it} \qquad (4.5.1.)$$

where MC_{it} and AC_{it} are the marginal and average costs of firm i at time t, L_{it} its labour input per unit of output and w_t the wagw rate, both at time t.

Following section 7 of the previous chapter, the learning curve may take one of two forms: $L_{it} = At^{-g}$ (4.5.2a.)

$$\text{or} \quad L_{it} = Ae^{-gt} \qquad (4.5.2b.)$$

Now , if all firms are <u>average cost pricers</u>:

$$p_{INit} = (1 + M_i)\ AC_{it} \qquad (4.5.3.)$$

where p_{INit} is the price of one ton of the innovation as supplied by firm i at time t and M_i is i's mark-up. Thus, the time path for price is given by:

$$p_{INit} = (1 + M_i)\ w_t\ At^{-g} \qquad (4.5.4a.)$$

$$\text{or} \quad p_{INit} = (1 + M_i)\ w_t\ Ae^{-gt} \qquad (4.5.4b.)$$

If all firms are marginal cost pricers:

$$p_{INit} = \frac{\eta_i}{\eta_i - 1}\ MC_{it} \qquad (4.5.5.)$$

1. This may not violate the real world too much. Stoneman, (1974, op.cit.), finds that the computer industry, for instance, by technological necessity, uses very labour-intensive methods.

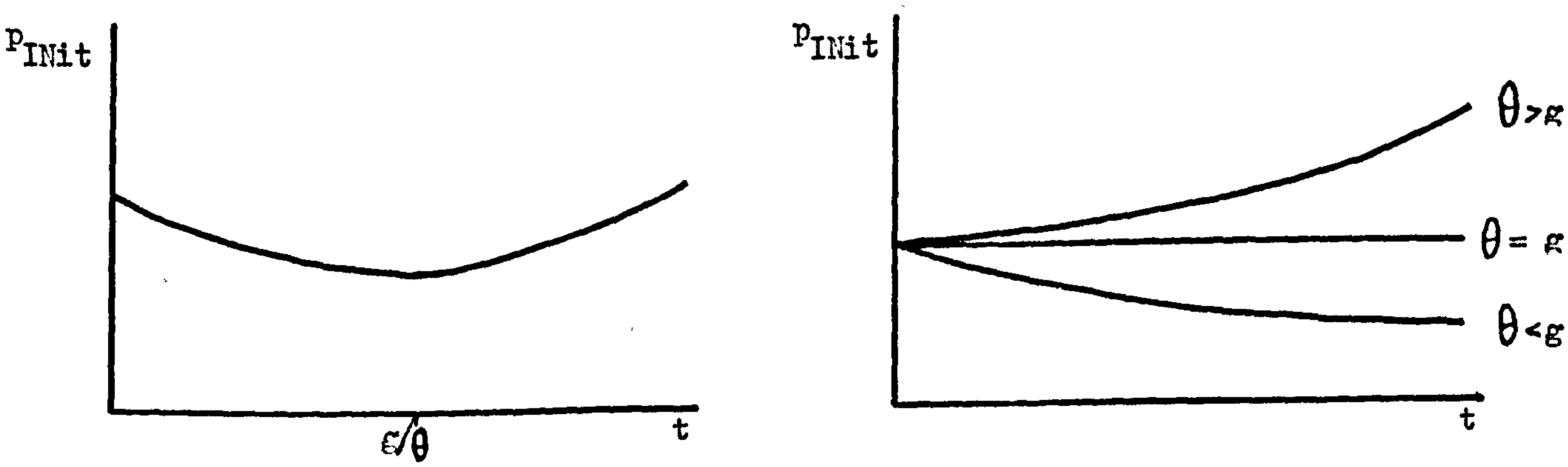
$$p_{INit} = \left(\frac{\eta_i}{\eta_i - 1}\right) w_t\, At^{-g} \qquad (4.5.6a)$$

$$\text{or} = \left(\frac{\eta_i}{\eta_i - 1}\right) w_t A e^{-gt} \qquad (4.5.6b)$$

where $-\eta_i$ is firm i's own price elasticity of demand.[1] Assuming $(1 + M_i)$ or $\left(\frac{\eta_i}{\eta_i - 1}\right)$ to be constant over time, and a growth in wages at a constant rate θ, the growth path of innovation price is as shown in figure (4.5.1).

<u>Figure 4.5.1. The growth path of innovation price.</u>

Group A innovations(4.6.4a. or 4.6.6a) Group B innovations(4.6.4b.or 4.6.6b)



Of course this highly simplified model fails to take into account a number of plausible possibilities. For instance, in the early years, producers may be prepared to sell at below cost in order to establish their innovation. Alternatively, they may prefer to sell, initially, at a relatively high price until sufficient capacity has been created to meet the accelerating demand which might be expected. Similarly, if they are forced to produce at a point of diminishing returns, because of capacity constraints, price will be high initially, until capacity is expanded by enough to permit production on the flat portion of the cost curve.[2]

Another possibility is that the mark-up or elasticity will change over time. As diffusion proceeds, the market nears saturation and replacement buying may not occur in some cases for a number of years. At the same time,

1. Where marginal revenue = price $\left(\frac{\eta - 1}{\eta}\right)$.

2. Some manufacturers were approached, by letter, on the question of capacity constraints. In the handful of replies received to this question, they were

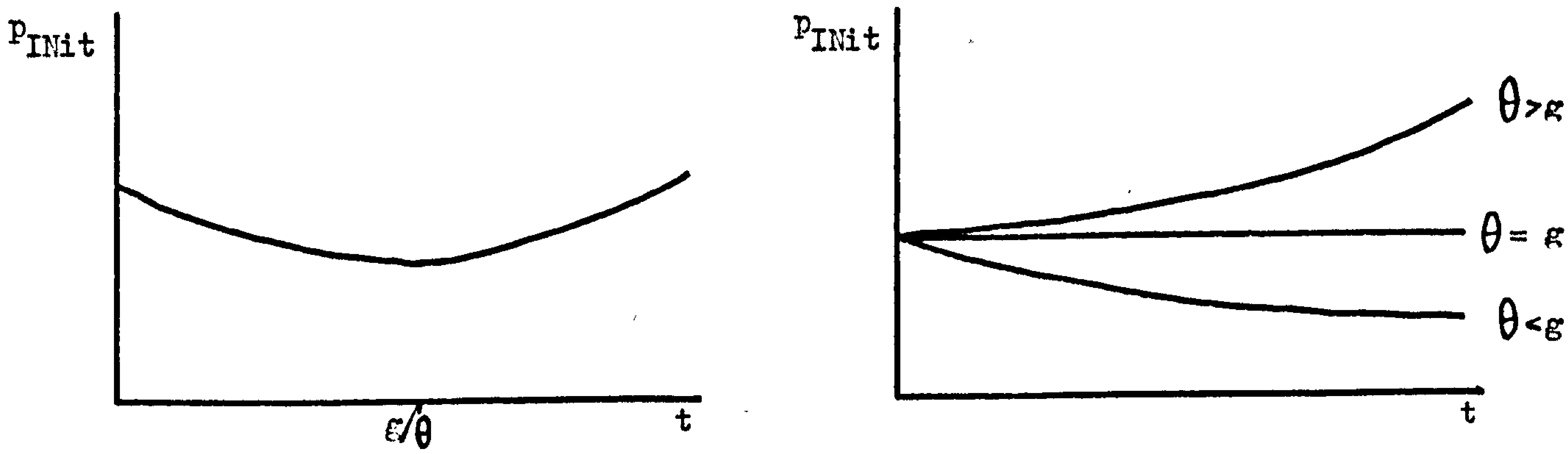
$$p_{INit} = \left(\frac{\eta_i}{\eta_i - 1}\right) w_t At^{-g} \qquad (4.5.6a)$$

$$\text{or} = \left(\frac{\eta_i}{\eta_i - 1}\right) w_t Ae^{-gt} \qquad (4.5.6b)$$

where $-\eta_i$ is firm i's own price elasticity of demand.[1] Assuming $(1 + M_i)$ or $\left(\frac{\eta_i}{\eta_i - 1}\right)$ to be constant over time, and a growth in wages at a constant rate θ, the growth path of innovation price is as shown in figure (4.5.1).

Figure 4.5.1. The growth path of innovation price.

Group A innovations(4.6.4a. or 4.6.6a) Group B innovations(4.6.4b.or 4.6.6b)



Of course this highly simplified model fails to take into account a number of plausible possibilities. For instance, in the early years, producers may be prepared to sell at below cost in order to establish their innovation. Alternatively, they may prefer to sell, initially, at a relatively high price until sufficient capacity has been created to meet the accelerating demand which might be expected. Similarly, if they are forced to produce at a point of diminishing returns, because of capacity constraints, price will be high initially, until capacity is expanded by enough to permit production on the flat portion of the cost curve.[2]

Another possibility is that the mark-up or elasticity will change over time. As diffusion proceeds, the market nears saturation and replacement buying may not occur in some cases for a number of years. At the same time,

1. Where marginal revenue = price $\left(\frac{\eta - 1}{\eta}\right)$.

2. Some manufacturers were approached, by letter, on the question of capacity constraints. In the handful of replies received to this question, they were claimed not to be a problem. Further, no mention of this possibility has

.. continued ...

suppliers will have built up capacity to satisfy the period of fast growth in demand. The producers may be thus forced into a price-cutting war as they fight over a declining market. In terms of equation (4.5.4), M_i may drift downwards over time and similarly η_i in (4.5.6.) This situation may be ameliorated to the extent that capacity can be switched to the production of newer innovations, of course.

On balance, therefore, it is likely that the downward trend in innovation price will be re-inforced by these factors, but this can only be a very tentative conclusion.

Although the state of competition in the suppliers' industry might affect the <u>level</u> of innovation price (through M_i or η_i), there is little to suggest that it will influence the <u>growth rate</u> of price. Perhaps the possible price-war in the later stages of diffusion may be more pronounced the more competitive is the industry. Probably more important, however, is the likelihood that learning will be more rapid the fewer sellers there are. Assuming learning is not transferable across producers,[1] any given output will clearly produce more learning, the fewer firms by which the output is produced.

Unfortunately, the whole area of pricing of new processes is largely uncharted empirically. As already mentioned, the little evidence available from appendix one is consistent with falling price, but Enos's research is the only known example of a time series large enough to permit the curve fitting necessary to test the hypotheses suggested above.

...footnote continued from previous page...
1. been found in the technical literature. One alternative manifestation of such constraints would be long order books; again no evidence of this has been found. Nevertheless, this question is re-examined in chapter 10, section 2.

1. Even this possibility is lessened by the international nature of many of the supplying industries. British firms should have access to the learning experienced by their foreign subsidiaries or parents. Moreover, technical journals may tend to diffuse the results of learning quickly in such research intensive industries.

(d) Other areas of competition.

The rate at which the innovation improves in quality[1] may also be sensitive to the market structure of the supplying industry. Because of the oligopolistic nature of many of these industries, competition may be more concentrated into quality rather than price differences; consequently quality improvements may be less pronounced in the few sample innovations sold by monopolists. However, as was argued above, learning may be more rapid the fewer producers there are.

Similarly, the quality of information flows may be higher the more suppliers there are and the more competitive they are. Potential consumers may feel they are receiving a better balanced view if they have access to more than one supplier. Furthermore, the problem (mentioned in section 2) of a limited pool of information-purveyors (i.e. salesmen) may be less acute the more suppliers there are. Crudely, the extent of supplier-consumer contact should be greater, the fewer consumers and the more suppliers there are, other things being equal - which, of course, they may not be.

Conclusions.

It was argued in chapter 2 that the major theoretical weakness of the epidemic models is their rather cursory analysis of the adoption decision at firm-level. The purpose of this chapter has been to provide a somewhat deeper analysis, partly on the basis of empirical findings in certain areas.

The evidence considered in section 1 suggests that the evaluation methods used in the adoption decision are often rudimentary and satisficing, rather than optimising. From the evidence considered in section 2 it is obvious that the diffusion of information is non-trivial. Whilst there are good reasons for supposing that most potential adopters 'hear of' new innovations fairly quickly, they may continue in a state of at least partial ignorance

1. As measured, for example, by the labour productivity attained in using the innovation.

for a number of years. It would appear that the suppliers of new innovations are a main source of information, particularly where effective competition in the consuming industry limits the extent of information interchange between potential adopters.[1]

Taken together the evidence cited in sections 1 and 2 seemsto be largely in accordance with the behavioural theory of the firm. Section 3 presents a brief formal statement of that theory, from which further implications emerge. Specifically, certain hypotheses are suggested as to the determinants of firms' attitudes to a new innovation. Similarly, some insights are gained into the reasons why these attitudes may change over time. Potential non-attainment of goals may lead to both increased search and a revision downwards of goals; the former reduces the uncertainty attached to non-adoption and the latter may be reflected in less demanding requirements of new investment projects. Both factors will tend to reduce the critical rate against which expected returns from adoption are judged. In turn, non-attainment of goals may be due to two types of change in the environment. Exogeneous pressures (not directly related to the innovation itself) will probably be cyclical, whilst endogeneous pressures (caused by other firms having adopted the innovation) will continually increase over time. Section 4 considers, in some depth, the proposition that the endogeneous pressures on non-adopters are directly related to the competitive structure of their industry. It is suggested that this proposition is indeed generally correct, whether marginal cost or average cost pricing is assumed. A less equivocal answer is possible to the question of whether or not incentives, or returns, to adopt are greater the more effective is competition. Finally, section 5 briefly describes the industries responsible for supplying the sample innovations. On the basis of this description a number of implications are drawn as to the influence on the adoption decision of the characteristics of the supplying industry.

1. This constitutes another reason for doubting a close conceptual link with the spread of diseases as is assumed in the epidemic models.

Chapter 5 : A Model of Diffusion.

The objective of the simple model[1] presented in this chapter is to provide a theoretical analysis of the diffusion of new process innovations, which is capable of predicting (a) the typical shape(s) of the diffusion growth curve and (b) the reasons for the parameters of that curve varying between industries and innovations. Thus, it may be seen as a direct competitor to the epidemic model as used by Mansfield and others.[2]

The descriptions and analysis of chapters 3 and 4 provide the basis for two conceptual improvements on the epidemic approach. It has been argued that the latter is deficient, in that it largely ignores the technical characteristics of process innovations and 'glosses over' the adoption decision at the individual firm level. It is hoped that by close reference to the findings of the previous two chapters, the model will, at least partially, avoid these deficiencies.

Section 1 contains a general mathematical statement of the adoption decision at the firm level in the form of six basic general equations. Section 2 suggests more specific forms for these equations and introduces three basic assumptions which simplify the ensuing algebra. Both these specific forms and assumptions are discussed in appendices at the end of the chapter. Section 3 examines the predicted relationship between firm size and the probability of having adopted the innovation - the so-called Quasi-Engel curve. Section 4 uses this cross-section prediction, with the added assumption that firm size is lognormally distributed, to generate the predicted shape of the diffusion growth curve. However, at this stage, diffusion is defined with respect to 'innovation-time'. Section 5 discusses the relationship between innovation time and calendar time; it is concluded that, as a special case, the two coincide and that the diffusion growth curve then takes on a symmetrical S shape similar to the logistic curve.

1. Based on Probit Analysis (see chapter 2, section 6.)
2. See chapter 2, sections 1 and 2.

Nevertheless, this will often not be the case and a number of other curves are equally likely. Up to this point, the model has been framed within the context of a given innovation diffusing in a given industry. Section 6 examines the ways in which the parameters of the model might vary between industries and innovations. At the end of the chapter, three appendices are added. The first two offer some support for the choice of specific mathematical forms and certain assumptions used in section 2, and the third develops the relationship (not tested in the ensuing empirical work) between the expected date of adoption for any firm and its size.

1. A probabilistic statement of the adoption decision.

In a study of inter-firm diffusion,[1] only the first adoption by each firm is considered. Thus the term 'adoption' implies only that the firm begins to produce <u>some</u> of its output using the new process innovation. For instance, a brick firm is defined as an adopter of Tunnel Kilns (TK) as soon as it installs <u>one</u> TK, even although part of its output may still be produced using old technology (Hoffman) kilns. Consequently, the adoption decision, in this sense, concerns what the potential adopter considers to be his most potentially profitable installation of the new process. To use the tunnel kiln example again, this might be the replacement of the oldest and least efficient Hoffman kiln. Of course, for certain innovations in certain firms (e.g. the Basic Oxygen Process (BOP) in small steel firms), adoption often implies a 100% change-over to the new process, but this is not always the case.

Throughout, it is assumed that the new innovation is embodied in new capital equipment supplied by a capital goods industry which has no vested interest in any firm in the consuming industry. Moreover, all potential adopters are assumed to 'know-of' the existence of the innovation once it is commercially operational. Thus, no potential adopter is barred from

1. Otherwise known as the imitation process.

adopting because of ignorance or patent restrictions - both assumptions seem fairly reasonable given the evidence of the previous chapters.[1]

Firms are assumed to use simple pay-back methods, based on undiscounted calculations, to evaluate their investment projects.[2] Thus by time t, firm i will have adopted, if its expectation of the pay-back period required (for the cost of the innovation to be recouped from increased net earnings) is less than, or equal to, some maximum critical period. This critical pay-back reflects the firms attitude[3] to the new innovation and will be determined by a number of variables (to be discussed presently).

Denoting the state of ownership (that is, having adopted the innovation), for firm i at time t, by $q_{it} = 1$ and the state of non-ownership by $q_{it} = 0$, then:

$$\begin{aligned} q_{it} &= 1 \quad \text{if } (ER)_{Nit} \leq R^*_{it} \\ q_{it} &= 0 \quad \text{if } (ER)_{Nit} > R^*_{it} \end{aligned} \qquad (5.1.1.)$$

Where $(ER)_{Nit}$ is the expected pay-back from adoption[4] and R^*_{it} is the critical maximum pay back which would be tolerated for adoption to be acceptable, both for firm i at time t.[5]

The probabilistic counterpart of these expressions is:

$$P\left\{q_{it} = 1\right\} = P\left\{(ER)_{Nit} \leq R^*_{it}\right\} \qquad (5.1.2.)$$

1. Chapter 3 section 2, Chapter 4 sections 2a and section 5. (In all remaining footnotes to this chapter, chapter will be abbreviated by 'ch' and section by 'sn'.)
2. See Ch.. 4, sn. 1.
3. See Ch. 4, sn 3(b).
4. Thus $(ER)_{Nit}$ is inversely related to the firm's assessment of the profitability of adoption. If factor prices and demand are constant over the pay-back period, then $(ER)_{Nit}$ is simply the reciprocal of the expected immediate rate of return.
5. At face value, this assumes away the problem of 'timing' i.e. of entrepreneurs not adopting when $(ER)_{Nit} \leq R^*_{it}$ in the hope of earning a better rate of return by waiting, say, for future improvements in the technology. This point may be partly covered by assuming that the level of R^*_{it} may be determined by expectations of future $(ER)_{Nit}$.

that is, the probability of ownership, at time t, is given by the probability that expected pay back will be shorter than the critical pay back period.

[Perhaps a slightly more realistic formulation for (5.1.1.) might be that i will have adopted by time t so long as at some time before t (or at t itself) the expected pay back was (is) shorter than the critical pay back period. In which case, (5.1.2.) would be replaced by:

$$P\left\{q_{it}=1\right\} = P\left\{(ER)_{Ni\tau} \leq R^{*}_{i\tau} \text{ for some } \tau \leq t\right\} \qquad (5.1.2a.)$$

Having said this, the specifications for ER_N and R^* to be presented in the remainder of this chapter suggest that (5.1.2.) will probably approximate (5.1.2a.) quite closely.[1] Consequently, (5.1.2.) will be retained in the following as development of the model soon becomes unmanageable if based on (5.1.2a

$(ER)_{Nit}$ is defined as the product of two components; the actual pay back which could be achieved by firm i (R_{Nit}) and the degree of ignorance of i about the capabilities of the new process, (H_{it}), both at time t:

$$(ER)_{Nit} = R_{Nit} \cdot H_{it} \qquad (5.1.3.)$$

It is tempting to assume that $H_{it} \geq 1$; that is, that firm i never over-estimates the profitability from adoption because it has imperfect information about the innovation. However, the possibility must be allowed that for some firms, imperfect knowledge causes over-optimism.

From the arguments of the previous two chapters, the following specifications are adopted:

$$R_{Nit} = f_1\,(\underset{-}{S_{it}}\,;\,\underset{?}{C_t}\,;\,\underset{-}{t}\,;\,\underset{?}{X_{ijt}} \text{ where } j=1,\ldots.r_1.) \qquad (5.1.4.)$$

$$H_{it} = f_2\,(\underset{?}{S_{it}}\,;\,\underset{?}{C_t}\,;\,\underset{-}{t}\,;\,\underset{?}{Y_{ijt}} \text{ where } j=1,\ldots.r_2.) \qquad (5.1.5.)$$

$$R^*_{it} = f_3\,(\underset{?}{S_{it}}\,;\,\underset{?}{C_t}\,;\,\underset{+}{t}\,;\,\underset{?}{Z_{ijt}} \text{ where } j=1,\ldots.r_3.) \qquad (5.1.6.)$$

where S_{it} is the size of firm i at time t, C_t is the stage in the business cycle at time t [2], t is the number of years after the first commercial appearance of the innovation and X_{ijt}, Y_{ijt} and Z_{ijt} a number of other, unspecified, characteristics of firm i at time t. Expected signs of first order partial derivatives are shown under each variable.

1. For a more detailed discussion, based on the estimates of later chapters, see Appendix 4 to this chapter.

2. A number of empirical proxies might be used for C, including industry.....

footnote 2. continued...
unemployment, the rate of capacity usage, investment expenditures etc. As a matter of convenience, C_t is defined as having high values at or near the peak of the cycle and low values at or near the trough.

At this stage, only inter-firm differences are considered. Obviously, the characteristics of the industry and innovation will also influence R_{Nit}, H_{it} and R^*_{it} but discussion of this is postponed until section 6, when the model is broadened to incorporate innovation and industry differences.

Although previous chapters provide the rationale for the inclusion of these variables and their expected derivatives, a short recapitulation is perhaps necessary.

(a) The actual pay-back, R_{Nit}

Six reasons were offered in chapter 3 for new process innovations being more profitable (and having lower R_N), ceterus paribus, for large firms.[1]

(i) Often, they exhibit economies of scale in both operating costs and investment costs.

(ii) Where this is not the case, economies of large numbers often occur.

(iii) Because of their continuous nature and/or high fixed costs, most new processes require high capacity usage to produce significant savings over the older technologies. Larger firms may be better able to achieve this, because of their greater flexibility: downturns in demand can be accommodated by slowing down usage of other, older, technology processes whilst maintaining capacity output from the new process. Small firms may not have sufficient demand or resources to retain sufficient old capacity to cushion the new process in this way.

(iv) Large firms are, perhaps, more likely to employ staff and management with sufficient skills to operate the often complex new technologies efficiently.

(v) The disruption often caused by adoption may be easier to absorb for large firms.

1. Ch. 3, sn 6.
2. Ch. 3, sn 8.

(vi) Large firms, just because of their size, are more likely to contain the ideal technical conditions for adoption.

Evidence was also provided to suggest that the potential profitability of new innovations will depend on a number of technical characteristics of the firms involved.[1] Particularly relevant are the nature of a firm's product, existing equipment and inputs. $\sum X_{ijt}$ in equation (5.1.4) represent such characteristics. Clearly, their effect will vary from innovation to innovation, but, without detailed information for each firm in the 22 sample industries, there is little chance of being able to measure these characteristics. Nevertheless, they may account for significant inter-firm differences.[2]

The stage of the trade cycle (C_t) might also influence the returns from adoption in two ways. At or approaching the peak of the trade cycle there is a greater probability of being able to operate the new process at full capacity over the pay back period. Given the nature of the short-run average cost curve for many new innovations[3] this may be essential for adoption to be profitable. On the other hand, the disruption often caused by adoption will be less costly in periods of slack demand. Thus the overall influence of C_t on R_{Nit} is uncertain and will probably vary from innovation to innovation.

It was also argued earlier that the new innovation, typically, becomes more profitable over time. Learning by doing by its manufacturer produces improved later vintages, not only in reduced variable input (notably labour) requirements, but also through a reduced per unit price relative to the cost of labour.[4] Much of the discussion was concerned with the different types of learning curve that might be expected; this aspect will be considered in section 5 of this chapter, at present all that is assumed is that the actual pay back declines over time.

1. Ch 3, sn 5.

2. An empirical assessment is made in Ch 7, sn 2 of the cost (in loss of explanatory power) of not being able to specify these and other variables more precisely.

3. Ch 3, sn 8.

4. Ch 3, sn 7 and Ch 4, sn 6. Appendix 1 to this chapter provides an example of how these two factors work together to influence R_N.

It is possible that the returns from adoption might be increasingly depressed over time (leading to a rising payback) through the effects of competition. However, the analysis of section 5 of the last chapter provided no justification for this hypothesis. Certainly, early adoption may produce super-normal profits, whilst late adoption merely transforms a position of loss making into one of normal profits, but it is not clear whether early adoption will actually produce a greater increase in net revenue.

(b) The level of ignorance, H_{it}.

Equation (5.1.5) suggests that the quality of information that a firm has about a new innovation will be determined by a number of its characteristics and that this information will improve over time and vary across the trade cycle.

It was argued[1] that larger firms are likely to possess better information because (i) on average, they might have larger engineering departments and better technically educated managers.
(ii) suppliers are likely to divert greater sales efforts to larger potential consumers and
(iii) they may have more extensive links with other firms in their industries. (In other words, not only are they more receptive, but they actually also receive more information.)

On the other hand, much of this information may be lost in the internal bureaucracies of large firms.

There are, of course, many other characteristics of firms (ΣY_{ijt}) which may be relevant, but no attempt will be made to specify them - as for ΣX_{ijt} above.

The quality of information changes with time because of search.[2] To a certain extent, an improving understanding over time is inevitable, given the advertising effort of suppliers and social intercourse with firms who

1. Ch. 4, sn. 2d.
2. Ch. 4, sns. 2b and 2c.

have already adopted. Probably more important however is active search which is often motivated by non-fulfillment of goals.[1] Pressures to search on non-adopters should build up over time because an increasing proportion of their competitors are achieving competitive advantages from adoption. In addition, pressures might vary across the trade cycle, but in what way is not certain. It might be expected, for instance, that, under boom conditions, most goals are being achieved (e.g. profits, sales etc.), but if significant weight is attached to avoidance of pressure on capacity and labour shortages, search may still increase.[2] The extent to which this search produces better information may depend on the nature of the innovation. Again, discussion of this point is delayed until section 5 of this chapter.

(c) The critical payback R^*_{it}

As has been argued, the critical payback used in the adoption decision may be thought of, crudely, as representing the firm's 'attitude' to the innovation. As such it may be determined by the firm's goals, the extent to which the innovation is seen as an aid to achievement of these goals, the relative bargaining power of the various factions in the management coalition and the amount of risk which adoption implies.

The discussion in the previous chapter[3] suggested a number of reasons for and against large firms having more favourable attitudes to new innovations. On the one hand, large firms may

(i) possess more scientifically trained managers who might be expected to argue for early adoption,

(ii) have easier access to the funds necessary for adoption of, sometimes, quite costly innovations and

(iii) view adoption with less apprehension as the consequences of failures are not proportionately so harmful.

1. Ch. 4, sns. 3c and 4.
2. Clearly, the position might vary between industries and innovations.
3. Ch. 4, sn. 3(b).

On the other hand, the bureaucracy which characterises many large firms may lead to conservative rules of thumb being used in the assessment of any risky project. Similarly, the fact that management is often not the function of the owners in many large corporations may mitigate against a favourable attitude towards new innovations. Whilst new innovations are usually very profitable,[1] they do often require extra effort from managers. In such cases, if the influence of the owners is weak, managers may use an extremely conservative critical pay back rate in evaluation.

Once again, it must be acknowledged that there are other characteristics of firms[2] (ΣZ_{ijt}) which may be important determinants of R^*_{it} but, for the moment, no attempt is made to specify them.

Within the behavioural framework, there are two reasons why R^*_{it} might vary with time: improved information resulting from search may reduce the risk premium required before adoption can take place and non-fulfillment of goals may lead to relaxation in yardsticks used in decision making.[3] Generally, the passing of time should lead to a reduction in R^*_{it}: the so-called endogeneous influences, caused by competitors having adopted the innovation, should both reduce risk and lead to non-fulfillment of goals. However, the more cyclical, exogeneous influences (represented by C_t) could lead to either upwards or downwards revisions in R^*_{it} [3]. As has been explained, the exact influence of the trade cycle is not clear: certain goals are easier to attain under boom conditions whilst others are easier to fulfill in times of slump.

2. A more specific form.

Equations (5.1.4) - (5.1.6) have so far been expressed as generally as is possible for the sake of simplicity. The more specific forms assumed in this section are, to a certain extent, arbitrary. In the

1. See, for instance, the typical paybacks quoted in appendix one.
2. Ch. 4, sn. 3b.
3. Ch. 4, sn. 3c.

first appendix to this chapter, however, a number of arguments are presented which suggest that the form chosen (linear-in-logs) may be the most appropriate.

Let
$$R_{Nit} = \alpha_1 S_{it}^{\beta_1} \theta_1 (t, c_t) \epsilon_{i1t} \qquad (5.2.1)$$

$$H_{it} = \alpha_2 S_{it}^{\beta_2} \theta_2 (t, c_t) \epsilon_{i2t} \qquad (5.2.2.)$$

$$R^*_{it} = \alpha_3 S_{it}^{\beta_3} \theta_3 (t, c_t) \epsilon_{i3t} \qquad (5.2.3.)$$

In other words, each of R_{Nit}, H_{it} and R^*_{it} are the product of four component parts. In (5.2.1), for instance, R_{Nit} is partly determined by a constant set of variables common to all firms (α_1) and a further set, again common to all firms, but which will vary over time and across the time cycle (θ_1). In addition, firm i's size also influences R_{Nit}, but with an elasticity common to all firms; the error term, ϵ_{i1t}, represents the influence of the other unspecified firm-level determinants, ΣX_{ijt}.

The precise mathematical forms of θ_1, θ_2 and θ_3 will be discussed in section 5, the only restrictions assumed at present being that

$$\frac{d\theta_1}{dt} < 0; \quad \frac{d\theta_2}{dt} < 0 \quad \text{and} \quad \frac{d\theta_3}{dt} > 0 \qquad (5.2.4.)$$

Now, from (5.1.3),

$$\frac{R^*_{it}}{(ER)_{Nit}} = \frac{R^*_{it}}{H_{it}.R_{Nit}}$$

which, using (5.2.1), (5.2.2) and (5.2.3),

$$= \frac{\alpha_3}{\alpha_1\alpha_2} S_{it}^{\beta_3 - \beta_2 - \beta_1} \cdot \frac{\theta_3}{\theta_1\theta_2} \cdot \frac{\epsilon_{i3t}}{\epsilon_{i2t}\cdot\epsilon_{i1t}} \qquad (5.2.5)$$

Then, defining

$$\beta_3 - \beta_2 - \beta_1 = \beta, \ \frac{\alpha_3}{\alpha_1\alpha_2} = \alpha, \ \frac{\epsilon_{i3t}}{\epsilon_{i2t}\cdot\epsilon_{i1t}} = \epsilon_{it}, \frac{\theta_3}{\theta_2\theta_1} = \theta(t, c_t) \qquad (5.2.6)$$

where $\frac{d\theta}{dt} > 0$

yields: $$\frac{R^*_{it}}{(ER_N)_{it}} = \alpha S_{it}^{\beta} \; \theta_t \; \epsilon_{it} \qquad (5.2.7)$$

Together with equation (5.1.2), this expression provides the basis for an analysis of diffusion, which is firmly based on individual firms' decision making behaviour. However, the transition to the industry-level diffusion curve may only be effected by facing up directly to the aggregation problem.[1] In the following two sections, (5.2.7) is developed, by specifying the distribution across firms of first ϵ_{it} and then S_{it}. In section 5 precise mathematical forms are specified for θ_t.

<u>3. The Quasi-Engel curve.</u>

The multiplicative form of the Central Limit theorem states that if a variable, V, is the product of a large number of independent variables, then log V will have a normal distribution. Thus, V will be lognormally distributed. From (5.2.6.), ϵ_{it} is the product of three separate error terms, each determined by a large number of other variables ($\Sigma X_{ij}, \Sigma Y_{ij}, \Sigma Z_{ij}$.) Furthermore, it seems likely that these variables determine the individual error terms in a multiplicative way.[2] It is also probable that the technical characteristics represented by ΣX_{ij} will be largely independent of the management characteristics represented by ΣZ_{ij}. Although one might expect that some of ΣZ_{ij} will be correlated with some of ΣY_{ij}, there should still be a large number of independent variables in these two groups, especially given the almost infinite number of influences on information gathering and decision yardsticks in the behavioural firm.

Not unreasonably, therefore, a lognormal distribution will be assumed for ϵ_{it}.

1. Past models derived from the theory of epidemics by-pass this problem because they ignore the determinants of the <u>individual</u> firm's behaviour, see sections 1 and 2 of chapter 2.

2. It seems unlikely that they will act in an additive manner. For instance, if X_{i1} is defined as the nature of firm i's product and X_{i2}, as the age of its existing machinery, one would expect that the influence on profitability of X_{i1} would be partly affected by the level of X_{i2} - this is certainly the case for SP, for example. This interdependence of influence should also apply to ΣY_{ij} and ΣZ_{ij}.

Using Aitchison and Brown's notation[1], let

$$\epsilon_{it} \text{ be } \Lambda(0, \sigma^2) \text{ and } \frac{d\sigma^2}{dt} = 0 \qquad (5.3.1.)$$

That is, the log of the error term is normally distributed, with mean 0 and variance σ^2, the latter being constant over time.

The assumption of a mean of zero for $\log \epsilon_{it}$, is merely for expositional convenience: it merely confers on to α the role of the geometric mean of $(R^*/ER_N)_{it}$, after allowing for the influence of firm size. The constant variance for the error term implies that each firm has a value for (R^*/ER_N) which is stable relative to all other firms', (abstracting from the differential changes in firm size.) This is not to say that this ratio remains constant over time, but only that it will change in the same way for all firms.[2]

Whilst these assumptions are not enough, on their own, to provide a prediction for the shape of the diffusion growth curve, they do have implications for the cross-section relationship between firm size and adoption behaviour.

Equation (5.1.2.) may be re-expressed as:

$$P\left\{q_{it} = 1\right\} = P\left\{\frac{(ER_N)_{it}}{R^*_{it}} \leq 1\right\} \qquad (5.3.2.)$$

and, from equation (5.2.7.),

$$P\left\{q_{it} = 1\right\} = P\left\{(\alpha S_{it}^{\beta} \; \theta_t \; \epsilon_i)^{-1} \leq 1\right\} \qquad (5.3.3.)$$

$$= P\left\{S_{it} \geq (\alpha \; \theta_t \; \epsilon_i)^{-1/\beta}\right\} \quad \text{for } \beta > 0 \qquad (5.3.4.)$$

The assumption that $\beta > 0$ is retained for most of this chapter. The cross-section empirical work of chapter 7 suggests, fairly conclusively,

1. op. cit., chapter 2. An alternative way of expressing (5.3.1.) is '$\log \epsilon_{it}$ is $N(0, \sigma^2)$'.
2. The implications of a variable σ^2 are considered in appendix 2 to this chapter.

that this is the case for all of the sample innovations.[1]

At this point, the exposition may be eased by the following definition:

$$S_{cit} = (\alpha\, \theta_t\, \epsilon_i)^{-1/\beta} \qquad (5.3.5.)$$

Then, the probability that firm i has adopted the innovation by time t may be re-expressed as the probability that its size is not less than some critical size, S_{cit}.[2]

Thus, $$P\left\{q_{it} = 1 \,\middle|\, S_{it}\right\} = P\left\{S_{cit} \leq S_{it}\right\} \qquad (5.3.6.)$$

Given that ϵ_i is $\Lambda(0, \sigma^2)$ and α and θ_t are constant across firms at time t, then S_{cit} will also be lognormally distributed:[3]

$$S_{cit} \text{ is } \Lambda\left(\frac{-\log\alpha - \log\theta_t}{\beta}, \frac{\sigma^2}{\beta^2}\right) \qquad (5.3.7.)$$

In which case,

$$P\left\{q_{it} = 1 \,\middle|\, S_{it}\right\} = \Lambda\left(S_{it} \,\middle|\, \frac{-\log\alpha - \log\theta_t}{\beta}, \frac{\sigma^2}{\beta^2}\right) \qquad (5.3.8.)$$

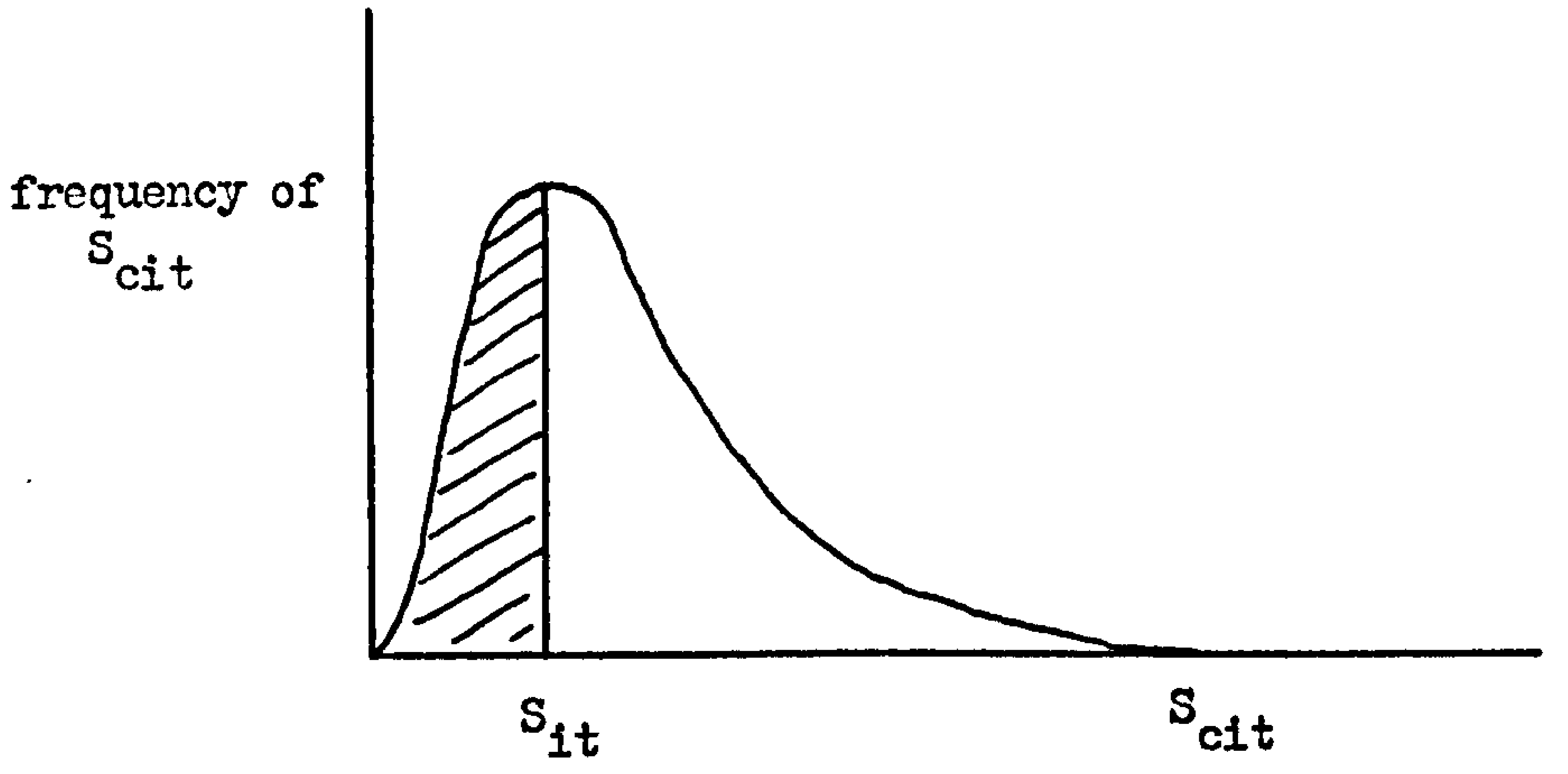
That is, the probability that firm i will have adopted by time t is equal to the proportion of the critical size distribution that lies to the left of the point S_{it}. In terms of figure 5.3.1., this probability is equal to the shaded area as a proportion of the total area under the curve.

1. It has been suggested earlier that economies of scale would ensure that $\beta_1 < 0$, but the signs of β_2 and β_3 have been left unspecified as there are arguments both ways (see section 1 of this chapter.) The weight of these arguments does tend to suggest a positive sign, overall, for β, but the matter is by no means certain. Under such circumstances, this question is best resolved empirically. It should be stressed that the model is capable of handling negative values for β, however.

2. It is not suggested that entrepreneurs see the adoption decision in this way however. The concept of a critical size is introduced purely as an expositional convenience.

3. See Aitchison and Brown's theorem 2.2., op.cit., p.11.

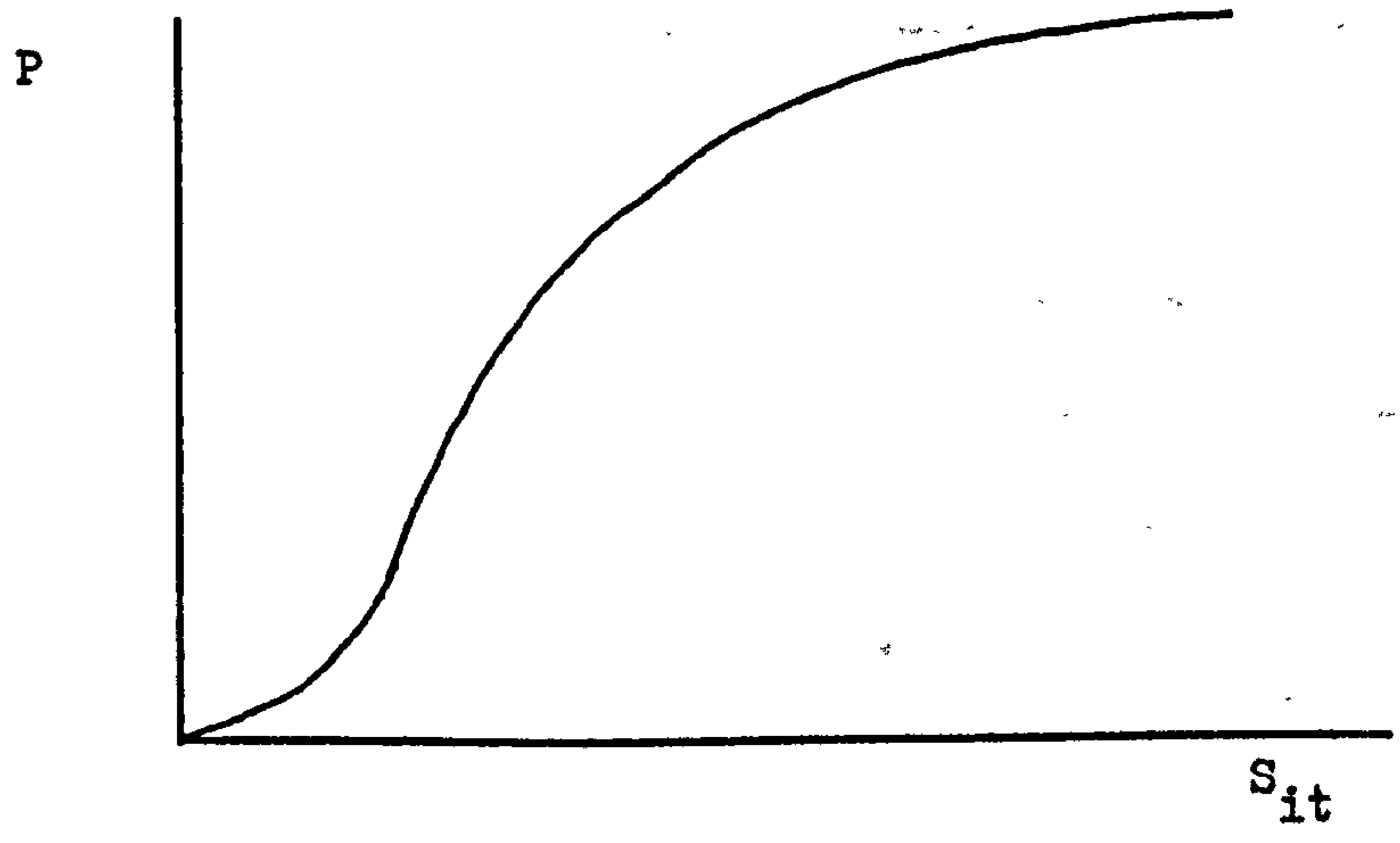
Figure 5.3.1. The critical size distribution.



Therefore, P rises monotonically with S_{it}. Specifically, the probability of having adopted, by time t, increases with firm size according to the cumulative lognormal Quasi-Engel curve of figure (5.3.2.) (This is analogous to the Quasi-Engel curve in past work on the diffusion of consumer durables, which relates probability of ownership to consumer incomes.[1] Because of the close similarity, the same term is retained in this context.)

This curve is, of course, the cumulative version of the distribution in the previous figure, with mean and variance as in equation (5.3.7.)

Figure 5.3.2. The Quasi-Engel curve.

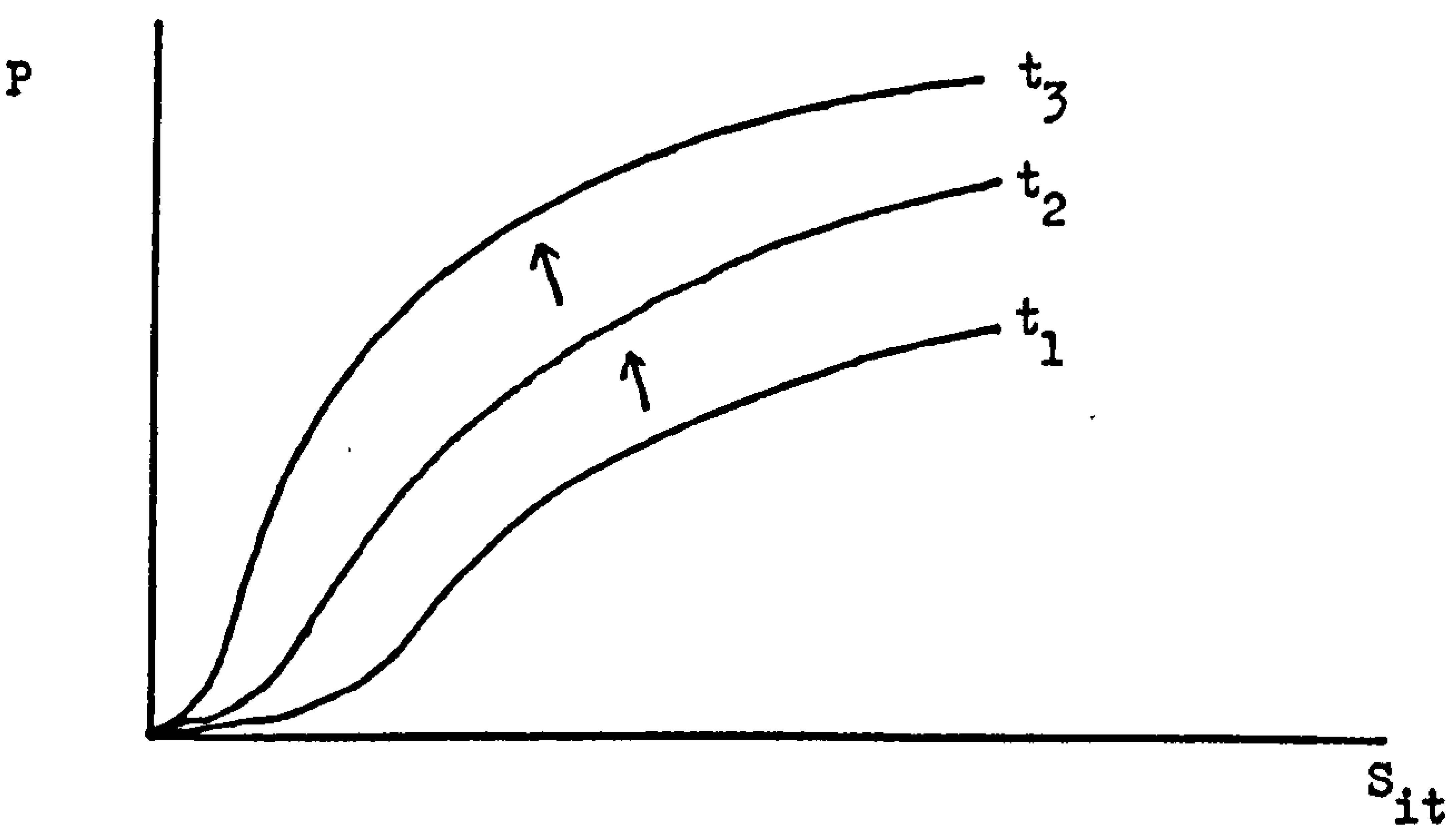


1. See chapter 2, section 6.

If, as has been assumed, β and σ^2 do not change over time, then the variance of this distribution remains constant over the diffusion period. The mean, however, will decline over time as θ_t becomes larger.[1] Thus the Quasi-Engel curve will always be cumulative lognormal, with a stable variance but a mean falling as diffusion proceeds. This change over time in the curve is portrayed in figure 5.3.3. The probability of having adopted rises for all firm sizes as the curve shifts continually to the left, and the point of inflexion approaches the origin.

Apart from its intrinsic interest value, the Quasi-Engel curve also provides a potential test of the above assumptions. Clearly the curve will only be cumulative lognormal if ϵ_i is lognormally distributed. An inspection of empirical Quasi-Engel curves should provide a test of this

Figure 5.3.3: The shift over time in the Quasi-Engel curve.



assumption therefore. Secondly, the relative importance of S_i and ϵ_i, as determinants of R^*_{it}, H_{it} and R_{Nit}, can be deduced indirectly from the variance of the Quasi-Engel curve, σ^2/β^2. If indeed, S_{it} is an important determinant, then σ^2 should be low relative to β. In chapter 7 these two tests will be applied to curves estimated for the sample innovations.

Finally, the implications of $\beta<0$ can be very briefly explained. If $\beta<0$, the inequality of (5.3.4) changes direction:

$$P\left\{q_{it}=1\right\} = P\left\{S_{it} \le (\alpha_t\,\theta_t\,\epsilon_i)^{-1/\beta}\right\} \qquad (5.3.4.a)$$

1. See equations (5.2.6.)

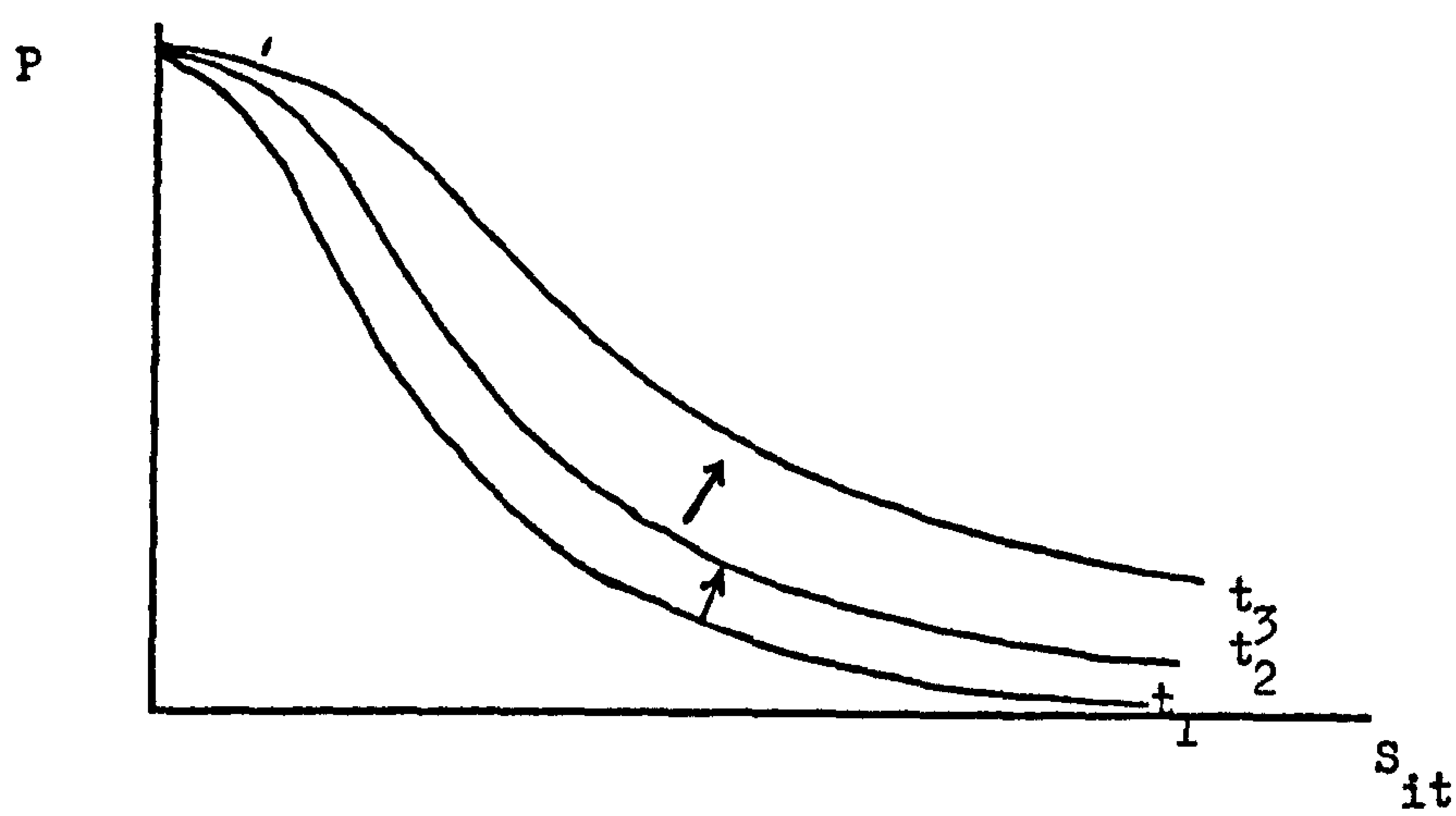
and correspondingly,

$$P\left\{q_{it}=1 \;\middle|\; S_{it}\right\} = P\left\{S_{cit} \geq S_{it}\right\} \qquad (5.3.6a)$$

$$= 1 - \Lambda\left(S_{it} \;\middle|\; \frac{-\log\alpha - \log\theta_t}{\beta}, \sigma^2/\beta^2\right) \qquad (5.3.8a)$$

Thus the Quasi-Engel curve is the complement of 5.3.8, as in figure 5.3.4.[1] In this case, the mean value of S_{cit} increases over time and the curve shifts to the right as diffusion proceeds. Again, the probability of having adopted increases for all S_{it} as time passes, but in this case, smaller firms always have a higher probability.

Figure 5.3.4. The Quasi-Engel curve for $\beta<0$.



4. The time series diffusion curve.

In this section, the Quasi-Engel curve is aggregated across the firm size distribution to produce an expression for the aggregate probability of adoption at time t. The time path of this expression is the probabilistic counterpart of the diffusion curve.

Let Q_t be the probability of any firm, taken at random, having adopted by time t, then

$$Q_t = P\left\{S_{ct} \leq S_t\right\} \qquad (5.4.1.)$$

Now, to proceed any further, the second part of the aggregation question must be resolved: the distribution of S_t must be formally specified. Once

1. And firm i only adopts when its critical size exceeds its actual size.

again, a lognormal distribution is assumed.

Thus S_t is $\Lambda(\mu_{st}, \sigma^2_{st})$ (5.4.2.)

Certainly, on the basis of the tests in Appendix 5, this appears to be a clearly acceptable approximation.[1]

Now, assuming that S_c and S are independent, and there is no reason why they should not be, their joint density may be written as the product of their marginal densities;[2] substituting in (5.3.7.) and (5.4.2.),

$$Q_t = \int_0^{\infty} \Lambda\left(S_t \,\middle|\, \frac{-\log\alpha - \log\theta_t}{\beta}, \frac{\sigma^2}{\beta^2}\right) d\Lambda\left(S_t \,|\, \mu_{st}, \sigma^2_{st}\right) \quad (5.4.3.)$$

By the convolution properties of normal distributions,[3] this may be written as:

$$Q_t = \Lambda\left(1 \,\middle|\, \frac{-\log\alpha - \log\theta_t}{\beta} - \mu_{st}, \sigma^2/\beta^2 + \sigma^2_{st}\right) \quad (5.4.4.)$$

In other words, (5.4.3.) is the sum of the probabilities of adoption at each firm size, weighted by the probability that that firm size will occur; that is, the Engel curve is aggregated across the firm size distribution.

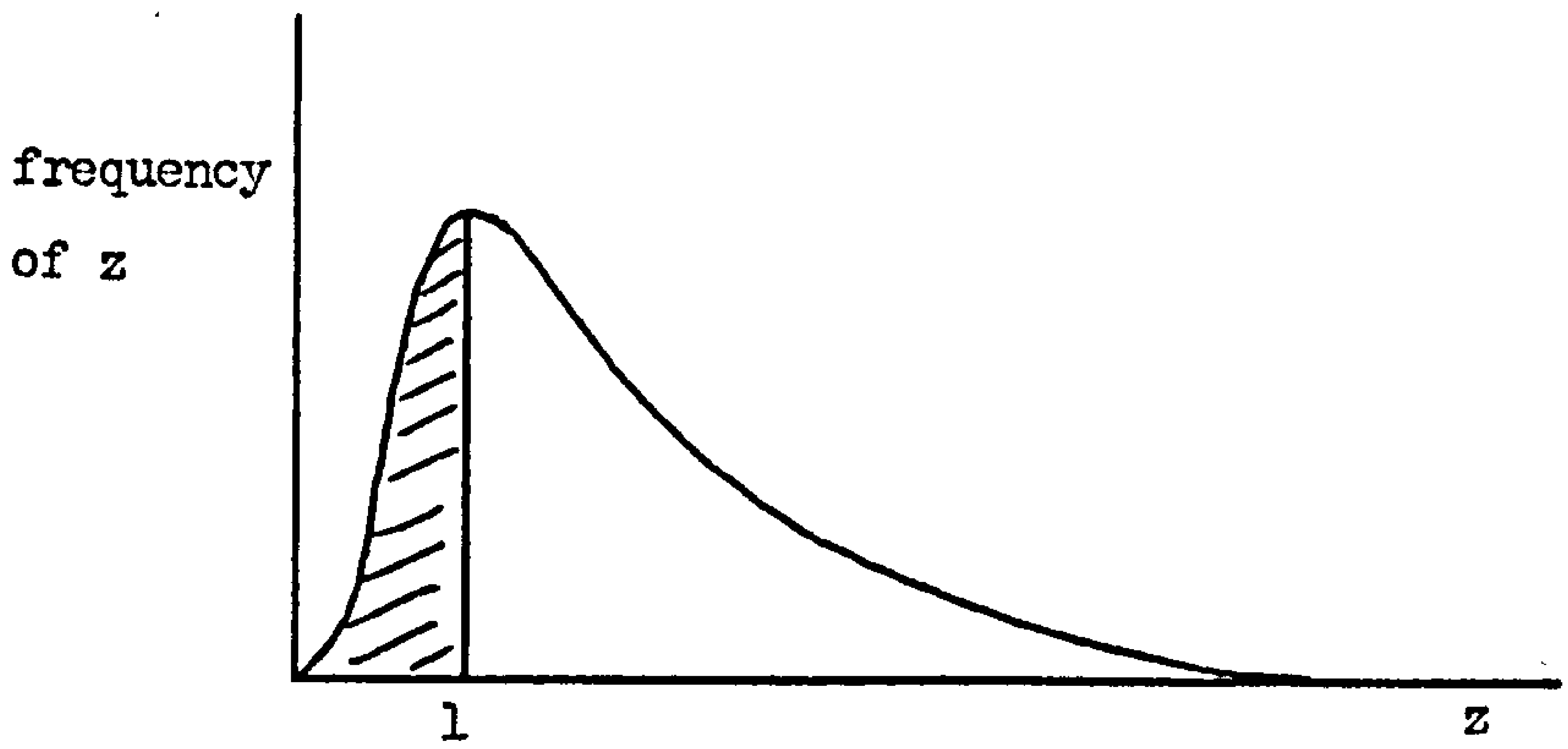
Because both S_{cit} and S_{it} are lognormally distributed, (5.4.3.) may be simplified as in (5.4.4.), which states that the aggregate probability of ownership of the innovation, at time t, is equal to the shaded area, as a proportion of the total area under the curve in figure 5.4.1.

1. See appendix 5. In only one industry was the data inconsistent with this assumption. The considerable literature on the applicability of the lognormal to size distributions is also surveyed in Appendix 5. Generally it has been found to be the most applicable of the usual distributions advocated.

2. See, for instance, Cramer, op.cit., p.37 and Aitchison and Brown, op. cit., p. 139.

3. ibid. p. 11.

Figure 5.4.1. A graphical representation of Q_t.



where z is $\Lambda\left(\dfrac{-\log\alpha - \log\theta_t \quad -\beta\mu_{st}}{\beta}, \quad \sigma^2/\beta^2 \quad + \sigma^2_{st}\right)$

From (5.4.4.), it follows that

$$Q_t \quad = \quad N\left(0 \;\middle|\; \frac{-\log\alpha - \log\theta_t}{\beta} - \mu_{st}, \sigma^2/\beta^2 + \sigma^2_{st}\right) \qquad (5.4.5.)$$

After manipulation,[1] this may be expressed as

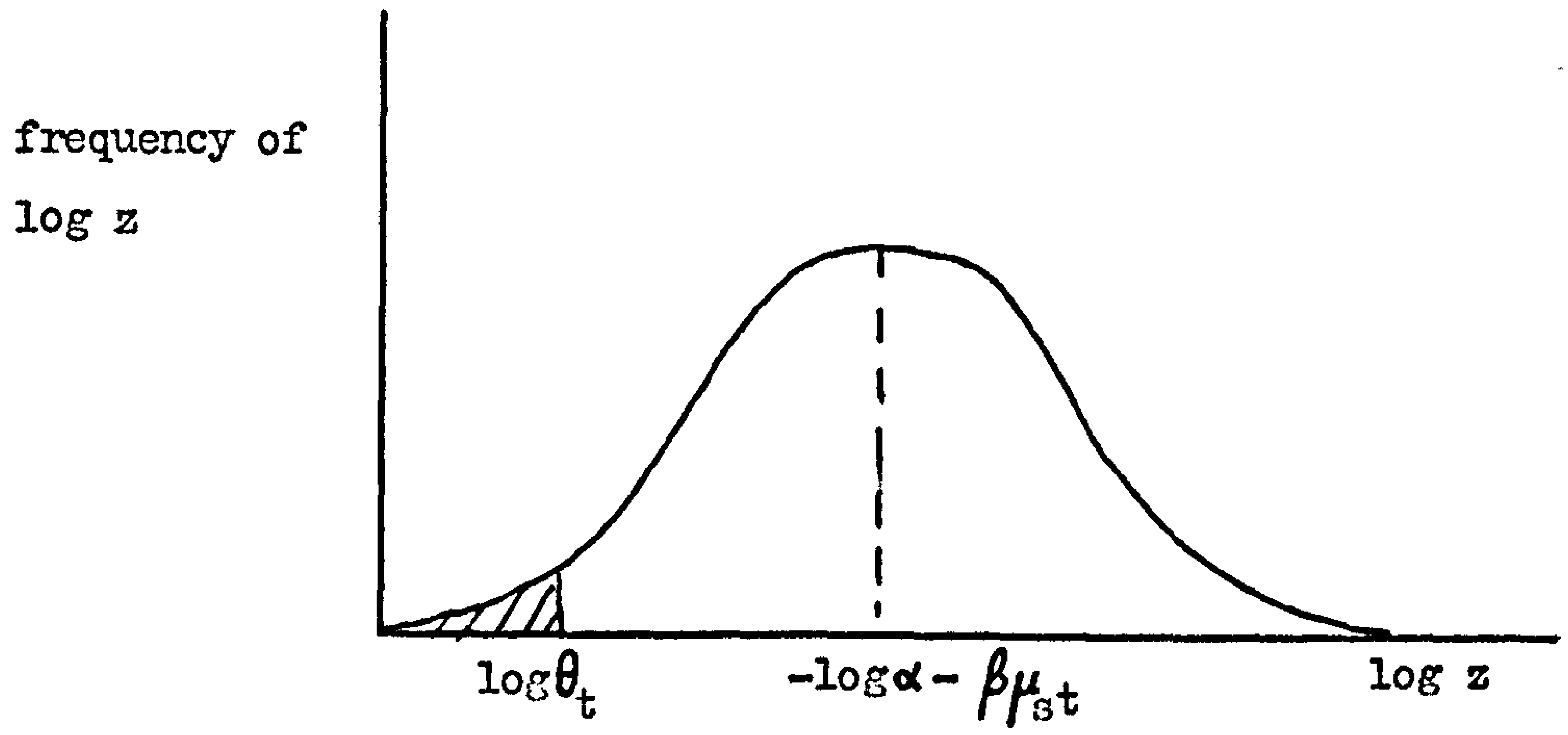
$$Q_t \quad = \quad N(\log\theta_t \mid -\log\alpha - \beta\mu_{st}, \sigma^2 + \beta^2\sigma^2_{st}) \qquad (5.4.6.)$$

Once again, the meaning of this expression is more clear from a diagram (figure 5.4.2.), in which Q_t is equal to the shaded area as a proportion of the total area under the curve.

1. From (5.4.5.), $Q_t = N\left(\dfrac{\log\alpha + \log\theta_t + \beta\mu_{st}}{\beta(\sigma^2/\beta^2 + \sigma^2_{st})^{1/2}} \;\middle|\; 0, 1\right)$

$$= N\left(\frac{\log\theta_t + (\log\alpha + \beta\mu_{st})}{(\sigma^2 + \beta^2\sigma^2_{st})^{1/2}} \;\middle|\; 0, 1\right) = N(\log\theta_t \mid -\log\alpha - \beta\mu_{st}, \sigma^2 + \beta^2\sigma^2_{st})$$

Figure 5.4.2. An alternative representation.

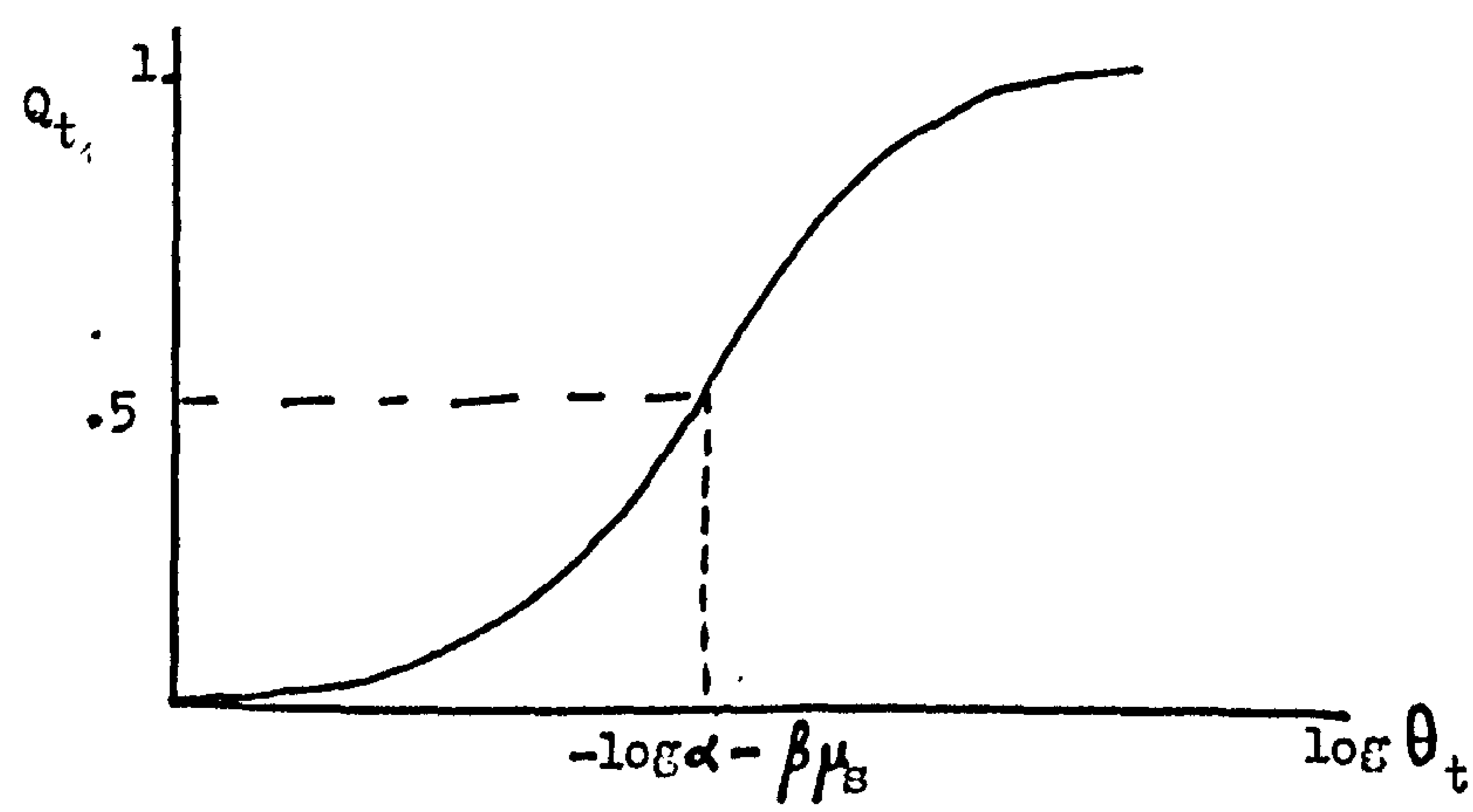


where log z is $N(-\log\alpha - \beta\mu_{st},\ \sigma^2 + \beta^2\sigma^2_{st})$

To derive the growth curve for Q_t, more information is needed. First, the time paths of μ_{st} and σ^2_{st} must be specified.

For the moment, μ_s and σ^2_{st} are assumed to be invariant with respect to time.[1] In which case, as $\log\alpha$, β and σ^2 are also constant, the mean and variance of the distribution (5.4.6.) must be stable over time. Therefore, the relationship between Q_t and $\log\theta_t$ is described by the cumulative form of a stable normal distribution, as shown in figure 5.4.3.

Figure 5.4.3. The aggregate growth curve plotted against $\log\theta_t$.



In other words, $Q_t = 0$

$\log\theta_t \to -\infty$

$Q_t = 1/2$ at $\log\theta_t = -\log\alpha - \beta\mu_s$

$Q_t = 1$ (5.4.7.)

$\log\theta_t \to \infty$

footnote 1. at top of next page...

1. There is little evidence to suggest that σ^2_{st} does change significantly over the time span of the typical diffusion process (say 20 years.) See, for instance, L. Engwall, 'Models of industrial structure,' Lexington Brooks (1973). In the only sample industry for which adequate data was available over a twenty year period, Lancashire weaving, σ^2_{st} changed by less than 5%. Both assumptions will be relaxed in Appendix 2 to this chapter. To anticipate its conclusions, a variable μ_{st} can be accommodated in the model quite easily, but whilst the implications of a variable σ^2_{st} are fairly clear intuitively, they are rather difficult to handle mathematically.

Thus, Q_t follows a symmetrical S shape (not dissimilar to a logistic[1]) with respect to $\log \theta_t$.

5. Innovation time: $\log \theta_t$.

One last step remains before the time series implications of the model emerge: $\log \theta_t$ must be specified, For the sake of convenience, $\log \theta_t$ is defined as "innovation time"; and the purpose of this section is, then, to specify the relationship between "innovation time" and conventional calendar time, t.

From (5.2.6.) $$\theta(t, c_t) = \frac{\theta_3(t, c_t)}{\theta_2(t,c_t)\,\theta_1(t,c_t)}$$

where $\frac{\partial\theta_1}{\partial t}<0$, $\frac{\partial\theta_2}{\partial t}<0$, $\frac{\partial\theta_3}{\partial t}>0$ and, therefore, $\frac{\partial\theta}{\partial t}>0$, but $\frac{\partial\theta}{\partial c_t} \gtrless 0$

In other words, θ_t increases with time, holding c_t constant, but the effects of c_t on θ_t, holding t constant, are more ambiguous.

(a) The influence of time.

The reasons for $\partial\theta/\partial t > 0$ may be briefly summarised as:

1) Improved specifications and reduced price, relative to labour costs, of later vintages of the innovation.

2) Improved information from passive search and active search motivated by endogeneous pressures[2] (e.g. increasing lack of competitiveness for non-

1. See chapter 2, equation (2.1.3a - 2.1.3c).
2. For a definition of endogeneous, as opposed to exogeneous pressures, see chapter 4, section 3c.

adopters.)

3) Reduced risk attached to adoption, due to improved information from the above-mentioned search.

4) Downward revisions in the critical rate (R*) because of non-fulfillment of goals due to endogeneous pressures.

Throughout the previous two chapters, it has been argued that the form of θ may vary depending on the type of innovation. Most of the discussion has concerned (1) in the above list, but similar implications must apply for (2) - (4). Ignoring, for the moment, the effect of C_t, two alternative forms that might be appropriate for θ_1 are:[1]

$$\theta_{1t} = t^{-\Psi_1} \qquad (5.5.1a)$$

or

$$\theta_{1t} = e^{-\Psi_1 t} \qquad (5.5.1b)$$

(given the existence of the constant intercept term, α, in (5.2.7), there is no necessity to include an initial value, θ_{10}, in these expressions.)

That is, the profitability of adoption increases over time (and thus the pay-back decreases) in both cases, but in (5.5.1a) with a decreasing growth rate, as opposed to a constant growth rate in (5.5.1b).[2] Retaining the nomenclature of chapter 3 section 7, (5.5.1a) might be expected to apply to group A type innovations - 'technically simple, probably relatively cheap and produced off-site,' whilst (5.5.1b) is more applicable to Group B 'more sophisticated, expensive innovations which are produced on a one-off basis, often requiring lengthy periods of installation on the adopter's site.'

By extension, furthermore, similar time paths might be expected for θ_2. For Group A innovations,

$$\theta_{2t} = t^{-\Psi_2} \qquad (5.5.2a)$$

and for Group B:

$$\theta_{2t} = e^{-\Psi_2 t} \qquad (5.5.2b)$$

1. Bearing in mind the conclusions of Ch.3, Sn. 7 and Ch. 4, sn. 6.
2. Appendix 1 to this chapter gives a precise example of how these forms might be derived from the arguments of the previous chapters.

that is, search will show diminishing returns much more readily for Group A[1].

Similarly, Group A innovations will be characterised by far less risk initially which will hardly decline, given the diminishing extra information resulting from search. In the same way, because of the declining rate of vintage-to-vintage improvements in Group A, the competitive pressures on non-adopters caused by the innovation are much less likely to accelerate as they might do for Group B innovations.[2]

Thus, again, for Group A:

$$\theta_{3t} = t^{\Psi_3} \qquad (5.5.3a)$$

and for Group B:

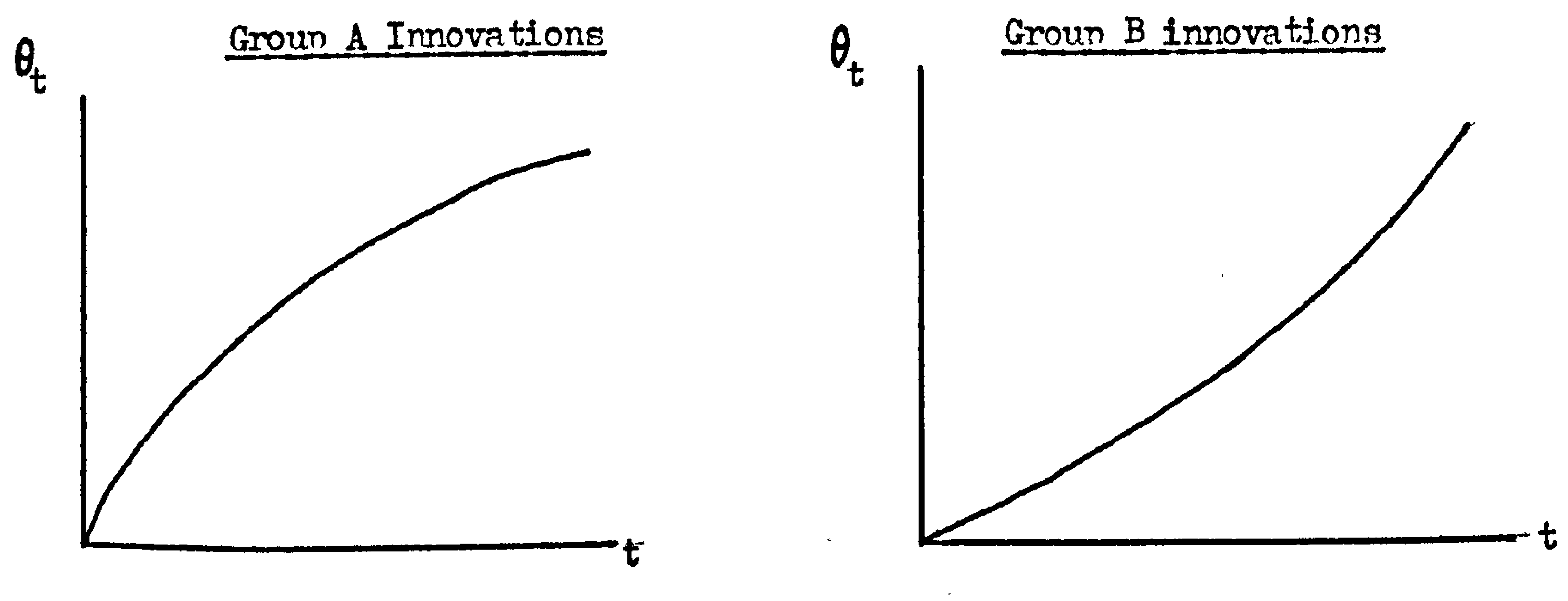
$$\theta_{3t} = e^{\Psi_3 t} \qquad (5.5.3b)$$

Combining (5.5.1 - 5.5.3) yields:

$$\theta_t = t^{(\Psi_3 + \Psi_2 + \Psi_1)} \quad \text{for Group A innovations} \qquad (5.5.4a)$$

and $$\theta_t = e^{(\Psi_1 + \Psi_2 + \Psi_3)t} \quad \text{" " B "} \qquad (5.5.4b)$$

Figure 5.5.1. Innovation versus calendar time.



1. See Ch., 4 sn. 2(d).
2. Ch., 4, sn. 4(iii).

(b) The influence of C_t.

At this point, at the loss of some generality, but with a compensating increase in clarity, C_t might be defined as the degree of capacity usage in the industry at time t[1] which, conceptually, might vary from 0 to 1.

The effects of C_t on θ_1, θ_2 and θ_3, holding t constant, can also be summarised briefly:

1) The profitability of the innovation, and the real cost of any disruption caused by its installation, will vary (in opposite ways) over the trade cycle.

2) Non-fulfillment of goals (and thus search) will also probably vary over the cycle and, in turn, this could effect:

(i) the rate of improvement in information about the innovation

(ii) the critical rate (R*) used to evaluate the innovation.

The net effect of C_t on θ_t is still unclear, however, for there is no way, other than empirically, of telling whether goals tend to be unfulfilled at the peak or the trough of the cycle.

Assuming for the moment that C_t has a net positive effect (namely that the profitability factor outweighs disruption and that non-fulfillment of goals and search activity is more probable at the peak of the cycle), a reasonable specification might be:

$$\theta_t = e^{\Omega C_t t^{\Psi}} \quad \text{for Group A} \qquad (5.5.5a)$$

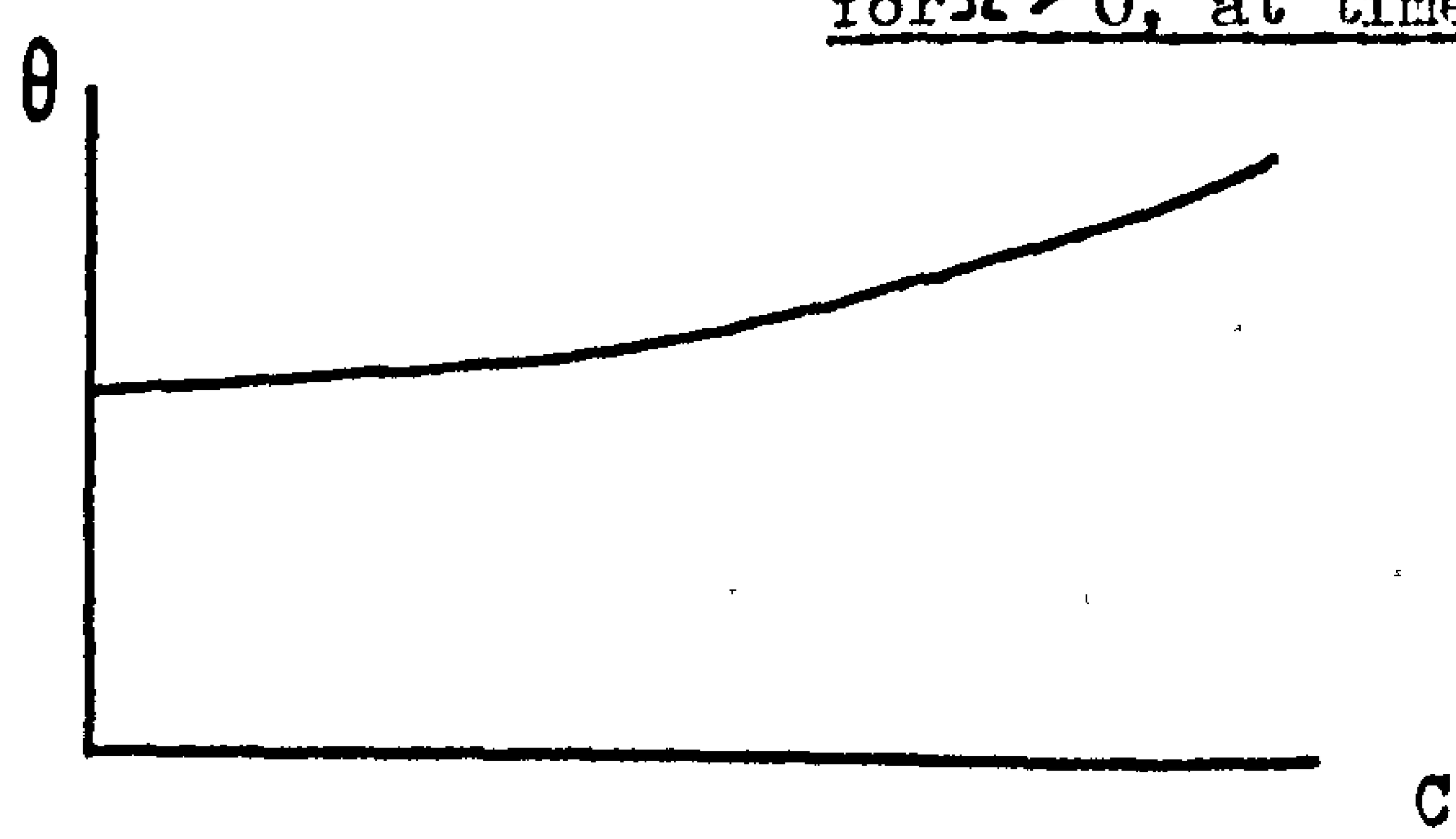
and

$$e^{\Omega C_t + \Psi t} \quad \text{for Group B} \qquad (5.5.5b)$$

where Ω is constant and positive.

Which has the graphical counterpart:

Figure 5.5.2. Innovation time versus capacity usage -(a) for $\Omega > 0$, at time t.

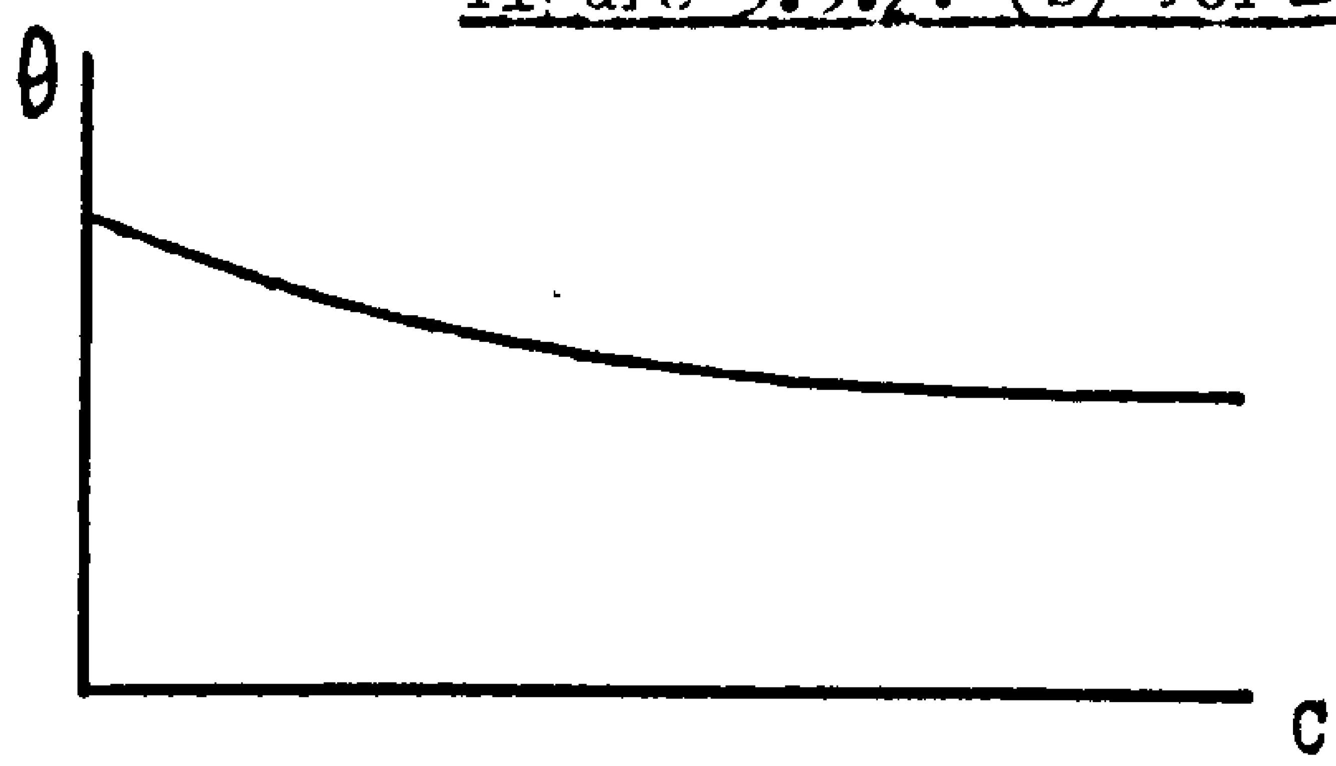


1. Of all empirical measures, capacity usage probably best reflects the state of the cycle. Nevertheless, it does only represent one aspect of the cycle; other aspects which might be relevant, but not perfectly correlated with capacity usage, are industry unemployment, profits, orders, investment etc. Hopefully all such variables will be sufficiently correlated with C_t, as defined, to justify its use as a proxy.

The implication is that θ increases <u>at an increasing rate</u> ($\partial^2\theta / \partial C^2 > 0$) the nearer the industry is to full capacity and the peak of the cycle.

Alternatively, if goals are unfulfilled more often at the trough of the cycle and the disruption effect outweighs the profitability effect, C_t has a net negative effect: $\Omega < 0$. This time, θ increases at an increasing rate, the nearer is the industry to the cycle trough.

Figure 5.5.2. (b) for $\Omega < 0$ at time t.



(c) Implications for the diffusion growth curve.

Given the simple forms for θ_t in equations (5.5.5), the implications follow very easily. Substituting into (5.4.6) the expression for log θ_t for a Group A innovation (5.5.5a):

$$Q_t = N\,(\psi \log t + \Omega C_t \mid -\log\alpha - \beta\mu_s,\ \sigma^2 + \beta^2\sigma_s^2) \quad (5.5.6a)$$

or for a group B innovation (5.5.b):

$$Q_t = N\,(\psi t + \Omega C_t \mid -\log\alpha - \beta\mu_s,\ \sigma^2 + \beta^2\sigma_s^2) \quad (5.5.6b)$$

Abstracting, for the moment, from the effects of C_t and setting $\Omega = 0$, these may be re-expressed:

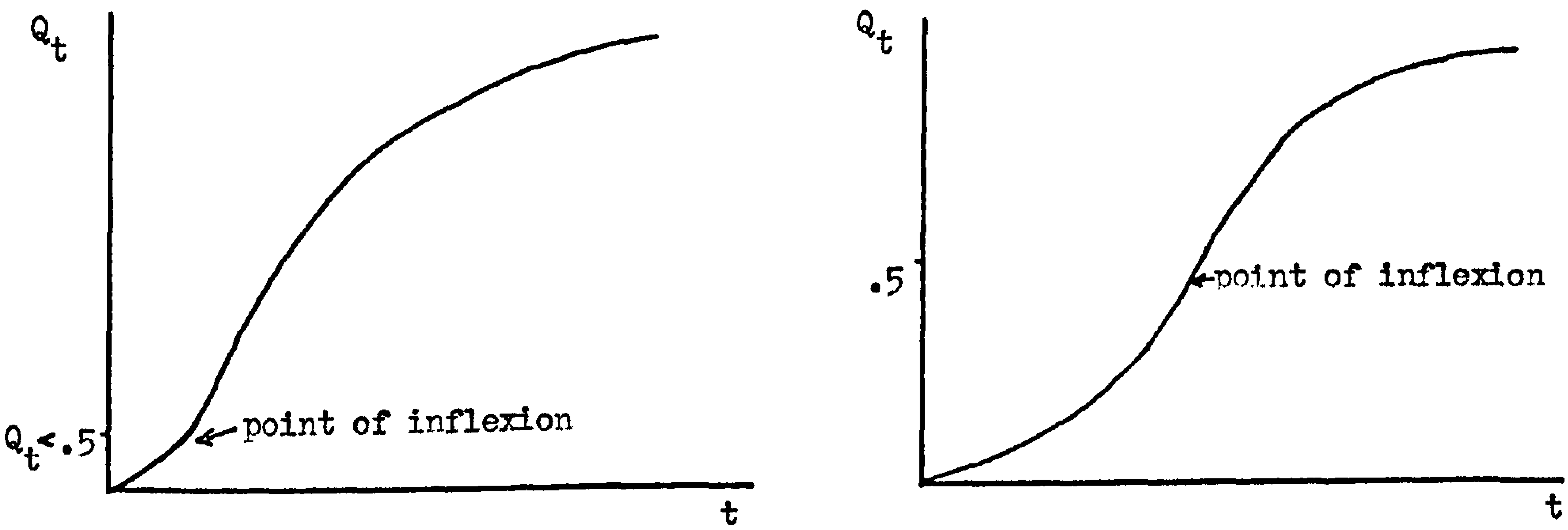
$$\left.\begin{aligned} Q_t &= N\left(\log t \,\middle|\, \frac{-\log\alpha - \beta\mu_s}{\psi},\ \frac{\sigma^2 + \beta^2\sigma_s^2}{\psi^2}\right) \\ &= \Lambda\left(t \,\middle|\, \frac{-\log\alpha - \beta\mu_s}{\psi},\ \frac{\sigma^2 + \beta^2\sigma_s^2}{\psi^2}\right) \end{aligned}\right\} \quad (5.5.7a)$$

and $$Q_t = N\left(t \,\middle|\, \frac{-\log\alpha - \beta\mu_s}{\psi},\ \frac{\sigma^2 + \beta^2\sigma_s^2}{\psi^2}\right) \qquad (5.5.7b.)$$

In other words, the growth curve of diffusion (as represented by Q_t), if plotted against calender time, t, will follow a cumulative lognormal curve for Group A innovations and a cumulative normal for Group B.

Figure 5.5.3. The aggregate growth curve abstracting from the influence of C_t.

(a) Group A innovations (b) Group B innovations



Re-introducing C_t into the argument involves only a minor complication. Given that C_t will exhibit a number of cycles over the diffusion period, but will have no pervasive underlying trend, the diffusion growth curve will also show cyclical variations around the underlying trends as shown in figure 5.5.3..

The mean and variance remain unchanged in (5.5.6.) but the variate does not increase at a uniform pace over time: rather, it slows down and speeds up according to the trade cycle.[1] Because C_t has strict upper and lower bounds,

1. It is just possible that innovation time could go backwards, given a sufficiently large positive Ω and a large negative change in C_t for example. This would require $\frac{dC}{dt} < -\frac{\psi}{\Omega t}$ (for 5.5.6a.) or $\frac{dC}{dt} < -\frac{\psi}{\Omega}$ (for 5.5.6b.). Estimates derived from the empirical diffusion curves suggest that these conditions are never satisfied for any of the sample innovations.

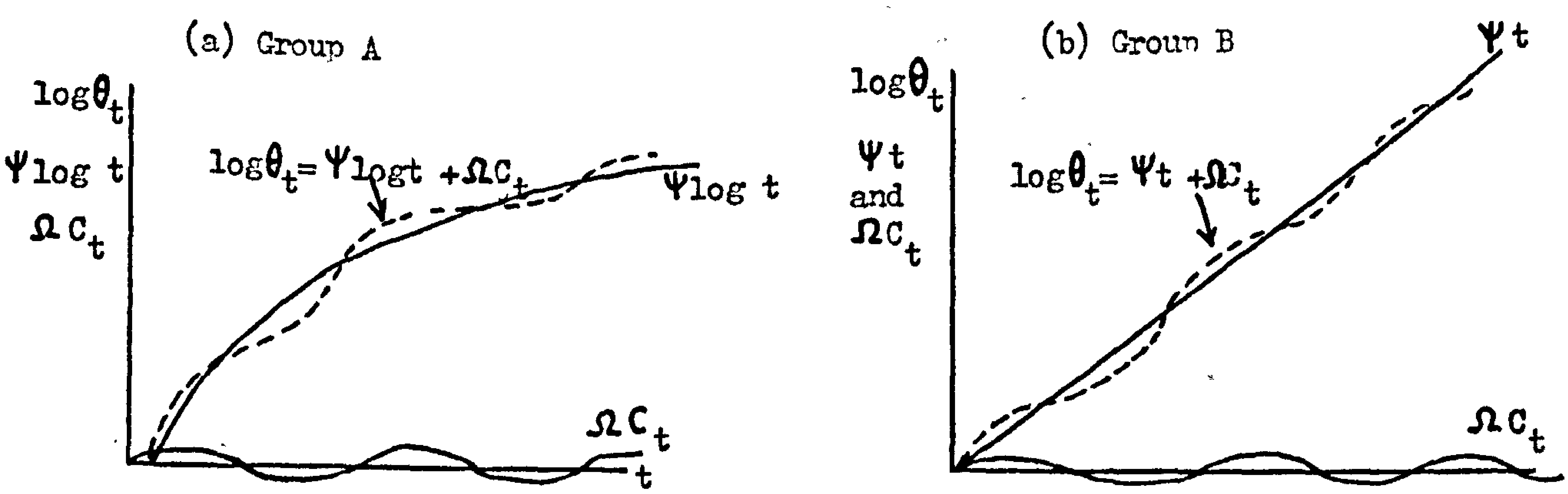
innovation time must, in the long run, approximate t or log t, but in the short-run it will only coincide at the mid-point in the trade cycle.

Graphically, the 4 broad alternative cases may be considered as in figure (5.5.5). Part (a) shows the diffusion curve for a Group A innovation with $\Omega \neq 0$. The curve in the second quadrant portrays the common factor in all 4 cases: namely that diffusion follows a symmetrical S shape if plotted against innovation time, $\log \theta_t$ (as in equation 5.4.6.) The curve in the third quadrant shows the relationship between innovation time and time ($\log \theta_t = \Psi \log t + \Omega C_t$). The fourth quadrant is empty, the 45° degree line simply serving to transfer measurements on the y axis to the x axis. Finally, the first quadrant shows the ensuing relationship between Q_t and time; the y ordinates are taken across from the fourth quadrant and the x ordinates, by a process of transformation, from the fourth via the third and second quadrants.

Similarly, part (b) constructs the diffusion curve with $\Omega = 0$ for Group A innovations and parts (c) and (d) perform similar functions for Group B innovations.

To ease the exposition of this already complicated diagram, the curve drawn in the third quadrant for parts (a) and (c), (that is, the relationship between t and $\log \theta_t$) were first constructed in a separate diagram (figure 5.5.4) assuming, as an example, a perfectly symmetrical cyclical behaviour for C_t.

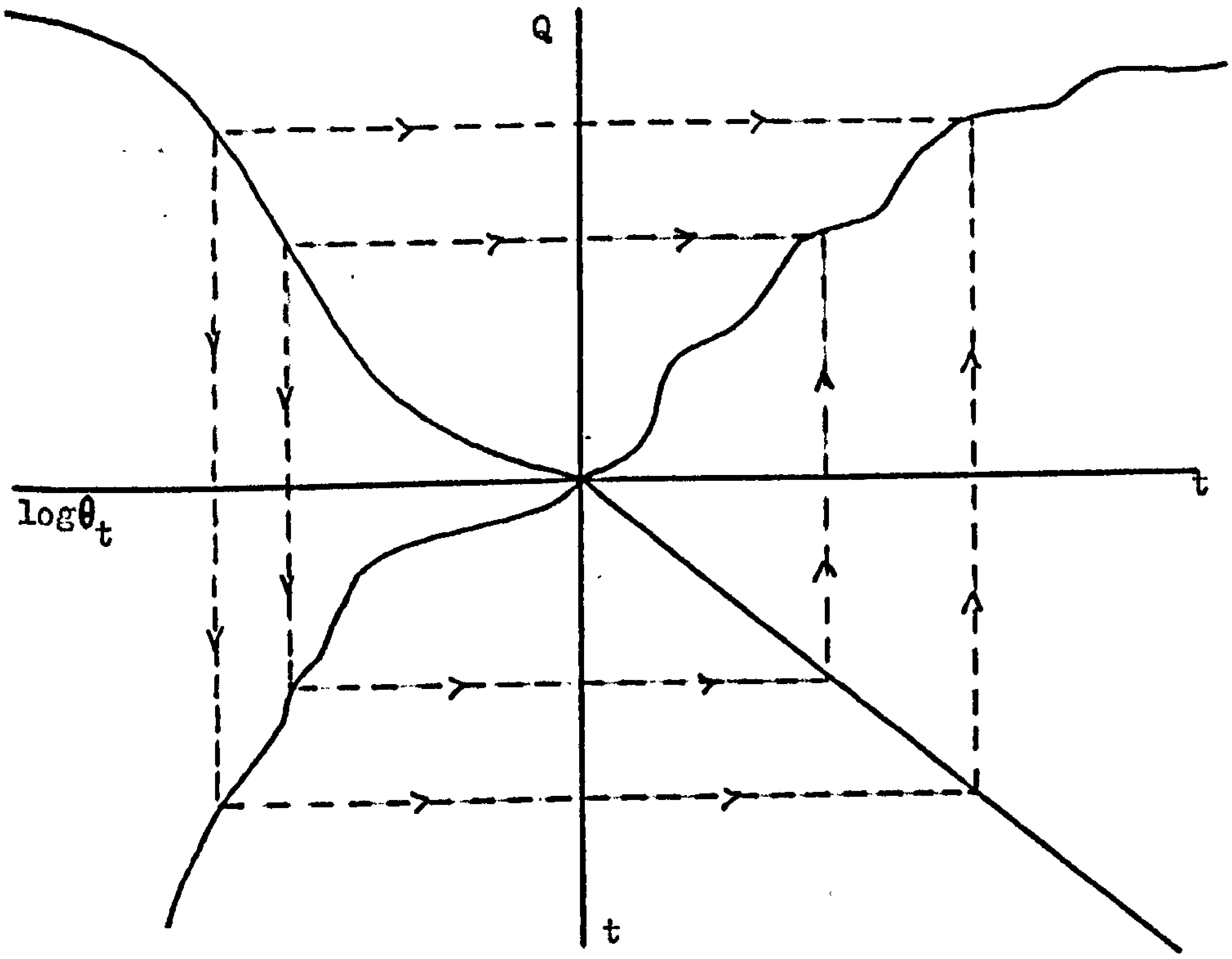
Figure 5.5.4. Innovation time versus calendar time including the influence of C_t



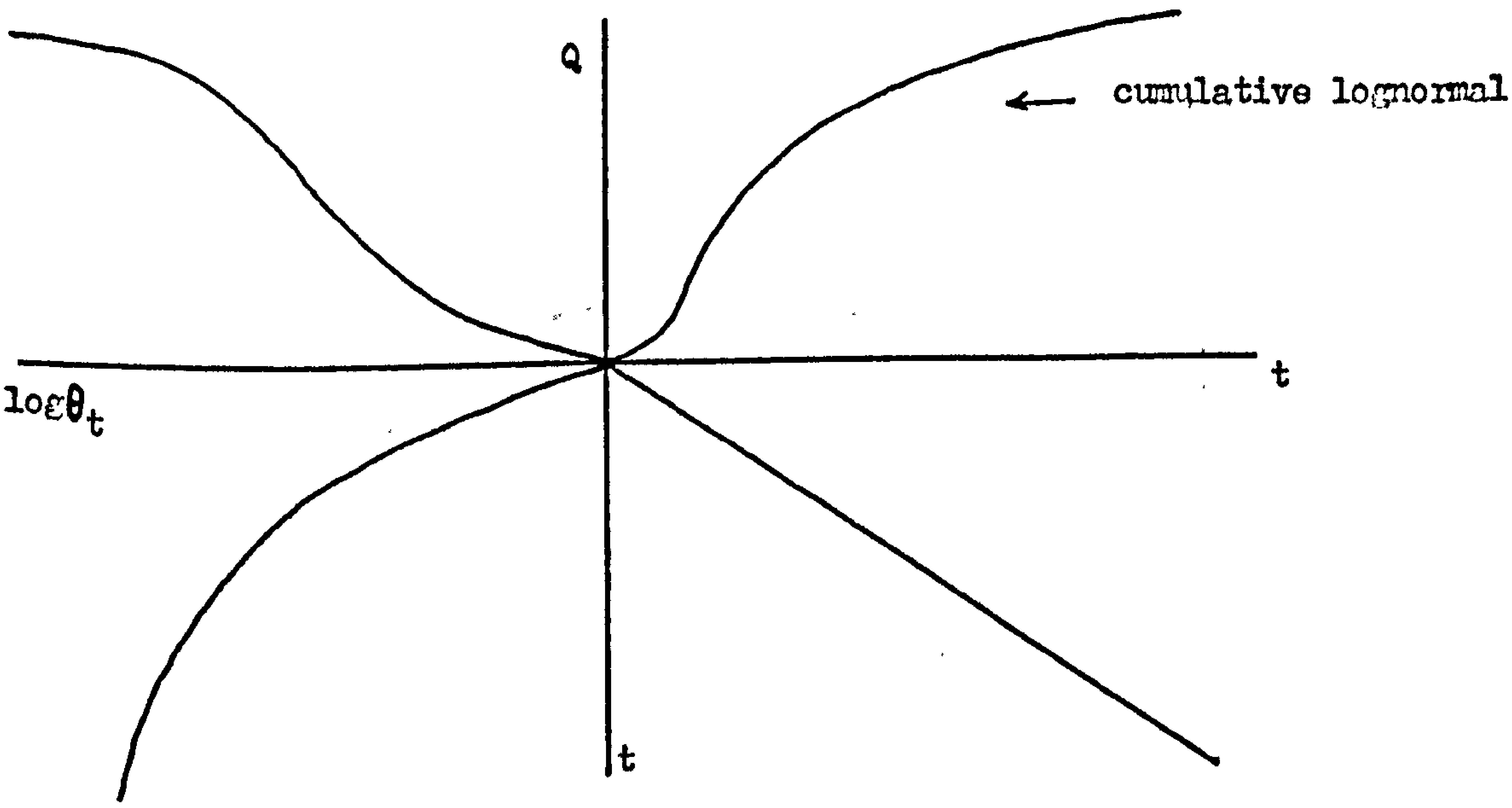
dotted line represents $\log \theta_t$

Figure 5.5.5: Diffusion plotted against "Innovation" and calendar time.

(a) Group A innovations with $\log \theta_t = \Psi \log t + \Omega C_t$

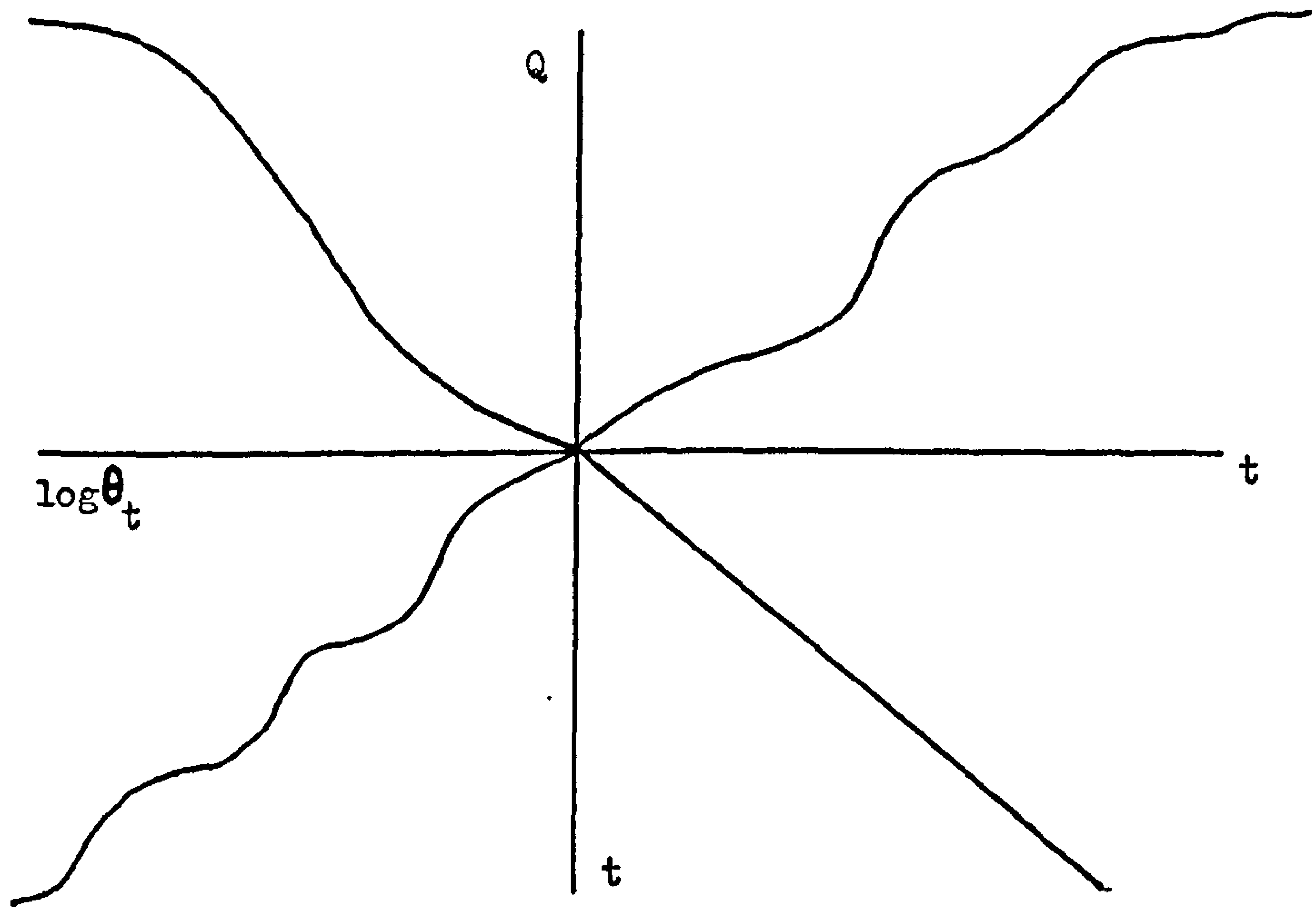


(b) Group A innovations with $\log \theta_t = \Psi \log t$

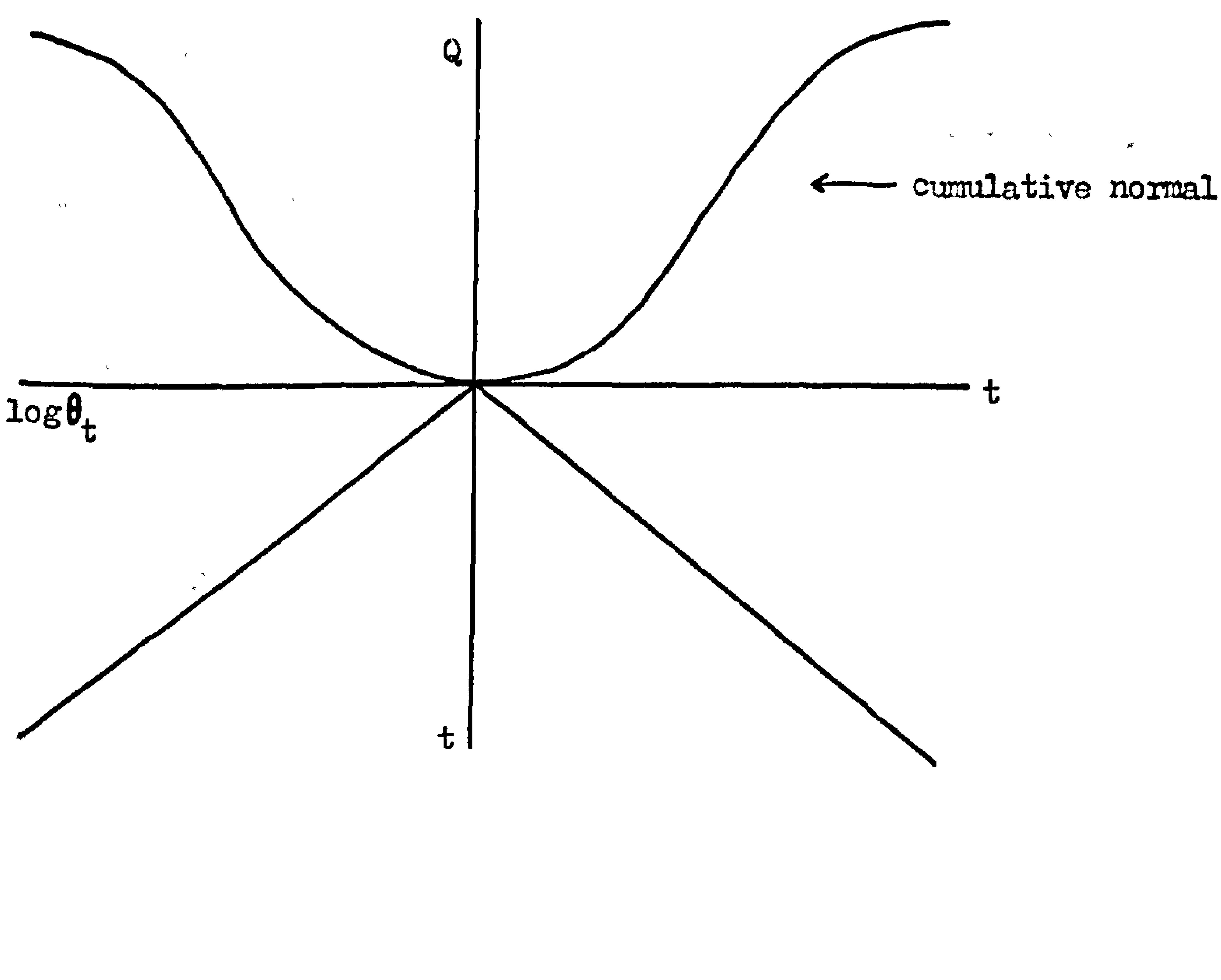


N.B. In order that $\log \theta_t$ may pass through the origin, C_o, for convenience, is set as equal to zero. In other words, C_t for all other $t \neq 0$ measures the deviation of capacity working in year t from the level in year 0. Thus, it will sometimes take on negative values.

5.5.5(c) Group B innovations with $\log \theta_t = \Psi t + \Omega C_t$



5.5.5(d) Group B innovations with $\log\theta_t = \Psi t$



Thus, there are four alternative basic forms for the diffusion growth curve. It is interesting to note that one of these (see figure 5.5.5.(d)), is a perfectly symmetrical S shape, which might be approximated by a logistic curve,[1] but this is only a special case: Group B innovations for which cyclical influences are unimportant.

The following chapter will be concerned with fitting these alternative curves to the time series diffusion data collected for the sample innovations.

6. Inter-industry and inter-innovation differences.

So far, this chapter has been concerned with the diffusion of a given innovation in a single industry. However, chapter 8 involves a cross-industry and innovation study of the determinants of the estimated parameters of these time series diffusion curves. From equations (5.5.6.), it can be seen that these parameters will be amalgams of the seven structural parameters of the model: α, β, Ψ, Ω, σ^2, σ_s^2 and μ_s. Thus, inter-industry differences will be the result of differences in these parameters.

This section paves the way for chapter 8 by drawing together the arguments of the previous chapters concerning the likely determinants of five of these structural parameters. An explanation of the other two, σ_s^2 and μ_s - the parameters of the size distributions, is outside the scope of the present study,[2] moreover, estimates are readily obtainable[3] for these parameters, unlike the other five.

(i) α_j

For an innovation diffusing in industry j, from (5.2.6),

$$\alpha_j = \frac{\alpha_{3j}}{\alpha_{1j}\alpha_{2j}}$$

1. But, as will become apparent in the next chapter, there are small, but sometimes crucial, differences between the cumulative normal and the logistic curve.

2. Involving, as it would, a theory of the determinants of industrial structure.

3. See Appendix 5.

Thus, it is a ratio based on the initial mean[1] values of R*, H and R_N. Crudely it will be determined by the initial mean pay-back (α_{1j}), the initial mean ignorance (α_{2j}) and the initial mean yardstick pay back (α_{3j}).

α_{1j} is simply related to the initial profitability of the innovation as claimed by its manufacturers' (π_{0j}) and, perhaps, the degree of competition in the consuming industry[2] (CC_j).

α_{2j} may also be determined by the degree of competition in the consuming industry[3] (CC_j), the number of firms in the consuming industry (N_j)[4], the typical profitability of the innovation, as claimed by the manufacturers (π_j)[4] and the technical complexity of the innovation (TI_j)[4].

α_{3j} should be influenced by competition in the consuming industry (CC_j), the initial outlay involved in adoption (K_j), the technical complexity of the innovation (TI_j) and the typical state of demand in the industry[5] (as represented perhaps, by the mean capacity-usage rate and the variance of that rate i.e. ($\bar{C}_j$) and ($V.C_j$).

Thus, let

$$\alpha_j = \varepsilon_1(\underset{+}{\pi_{0j}};\ \underset{+}{CC_j};\ \underset{-}{N_j};\ \underset{+}{\pi_j};\ \underset{-}{TI_j};\ \underset{-}{K_j};\ \underset{+}{\bar{C}_j};\ \underset{-}{VC_j}) \qquad (5.6.1)$$

The expected signs of the first order partial derivatives are shown under each variable.

<u>(ii) β_j</u>

From (5.2.6), $\beta_j = \beta_{3j} - \beta_{2j} - \beta_{1j}$.

As has been stated, the shapesof the estimated Quasi-Engel curves imply $\beta_j > 0$, for all j.

β_{1j}, β_{2j} and β_{3j} reflect, respectively, the extent to which firm size influences (a) the profitability of adoption, (b) the level of a firm's information and understanding of the innovation and (c) the firm's attitude towards the innovation.

1. After allowing for the influence of size.
2. See ch. 4, sn. 4.
3. See ch. 4, sn. 4.
4. See ch. 4, sn. 2(d)
5. See ch. 4, sn. 3b.
6. See ch. 7.

Generally, there is little in the discussion of the previous two chapters to suggest the probable determinants of β_j. Adams and Dirlam's evidence on the U.S. Steel industry[1] suggests that β_{3j} may be smaller, and possibly negative, when the industry concerned is not very competitive. (CC_j) On the other hand, scale economies are usually more pronounced for costly, technical, complex innovations (K_j and TI_j), particularly in the process industries. Similarly, large firms may be at a greater information advantage for such innovations.

Therefore, let

$$\beta_j = g_2 \left(\underset{+}{CC_j};\ \underset{+}{TI_j};\ \underset{+}{K_j} \right) \qquad (5.6.2)$$

<u>(iii) σ_j^2</u>

From (5.2.6), $\epsilon_{ij} = \dfrac{\epsilon_{i3j}}{\epsilon_{i2j}.\epsilon_{i1j}}$ Thus σ_j^2 should be the sum of the variances of each of the individual error terms.[2] As with β, previous chapters only provide a partial explanation of the determinants of each of these error terms.

The variance of ϵ_{i1j} should be larger the greater variability there is in the products sold, the inputs used and the age of the capital stock used by the industry concerned.[3] Generally, none of these are measurable, but one might expect greater homogeneity in product and inputs the more industrially and geographically concentrated is the industry (CC_j and GC_j respectively).[4]

1. See ch. 2, sn. 3b.
2. If the three error terms are independent, $\sigma = \sigma_1{}^2 + \sigma_2{}^2 + \sigma_3{}^2$; otherwise, covariance terms will also enter into the expression.
3. See ch. 3, sn. 5.
4. On the other hand, industrial concentration often leads to greater product differentiation, thus the effect of CC_j is by no means certain.

A proxy for the variability in age of existing capital stock is suggested by Salter. He argues that for industries in which new plant is capital intensive (and thus requiring high investment outlays), existing older equipment will have longer economic lives before being replaced:[1] 'It follows, therefore, that a wide range between the labour requirements of co-existing techniques is likely in highly mechanised industries. This is as we should expect; adjustment is most costly when highly mechanised techniques requiring large amounts of investment are involved.'[2] Thus the higher the labour intensity of the industry (LI_j), the lower might be the variance in age of existing machinery.

The variance of ϵ_{i2j} and ϵ_{i3j} are more problematic; one might expect both to be larger the more complex and costly the innovation (K_j and TI_j) and smaller the more science-based the consuming industry (SC). In addition, there may be less scope for widely divergent attitudes in highly competitive industries.[3]

Thus let

$$\sigma_j^2 = g_3\,(\underset{?}{CC_j};\ \underset{-}{GC_j};\ \underset{-}{LI_j};\ \underset{+}{TI_j};\ \underset{+}{K_j};\ \underset{-}{SC_j}) \qquad (5.6.3)$$

<u>(iv) Ψ_j</u>.

From equations (5.5.4), $\Psi_j = \Psi_{1j} + \Psi_{2j} + \Psi_{3j}$.

Ψ_{1j} reflects the [illegible] improvement over time in the profitability of the innovation. Therefore, from the discussions of learning by doing,[4] Ψ_{1j} should be higher, the more complex is the technology on which the innovation is based (TI_j). The effects of the market structure in the supplying industry

1. This follows from the usual vintage assumption that machinery is replaced only when its operating costs exceed the total costs of new equipment. Obviously, total costs (which include capital charges) will be larger, the more capital intensive the new equipment.

2. Salter, op.cit., p.71.

3. See ch. 4, sn. 3b.

4. Ch. 3, sn. 7 and Ch. 4, sn. 5.

(CS_j) are uncertain, but may be important. Two other characteristics of the supplying industry which may have an influence, but which are probably unquantifiable are the closeness of its historical links with the consuming industry[1] and how far it is internationally based.[1] It is also conceivable that the competitive structure of the consuming industry (CC_j) may determine Ψ_{1j} , but the analysis in chapter 4[2] could offer no hard conclusions on this matter.

The rate of improvement in firms' information about new innovations (Ψ_{2j}) will depend not only on the extent of search but also its yields. It has been argued that search will be relatively less productive (i.e. result in little improvement in information) the more competitive is the industry (CC_j)[3], the more firms there are in it (N_j)[3] and the simpler is the underlying technology (TT_j)[3] of the innovation. The <u>extent</u> of search may be positively correlated with the level of competition in the consuming industry (CC_j) as endogeneous pressures should be greater.[4] Similarly more profitable innovations (Π_j) should create more competitive pressures on non-adopters;[4] furthermore, they are more likely to be seen as a potential solution to non-attainment of goals.[5] These influences may be lessened somewhat by the reduced willingness of competitors to divulge information on profitable innovations[6] thus reducing the <u>yields</u> of search. Further, although exogeneous pressures have been envisaged as having a cyclical effect on search[7] (and thus determining Ω_j below), it is probable that they never subside altogether. Thus, the underlying non-cyclical component of these pressures will have a continuous effect on search, regardless of short-term fluctuations. For instance, an industry characterised by a higher <u>typical</u> level of capacity working ($\bar{C}_j$)

1. Ch. 4, sn. 5.
2. Ch. 4, sn. 4.
3. Ch. 4, sns. (b) and (c).
4. Ch. 4, sn. 4.
5. Ch. 4, sn. 3(c)
6. Ch. 4, sn. 2(b) and (c)
7. Ch. 4, sn. 3(b) and (c)

may be under consistently greater pressure to search if only because more investment opportunities will arise. The effect of fluctuations in C_{tj} may simply accentuate or lessen this underlying pressure.

Ψ_{3j} (the rate of improvement in attitudes to the new innovation) should be higher again where search is more intensive and productive if only because the risk premium is likely to be inversely related to the quality of information.[1] In addition, not only do pressures lead to greater information search, but also to a tendency to revise goals downwards. Thus, much the same set of variables should determine Ψ_{3j} as was postulated for Ψ_{2j}. Two other variables which might also affect the rate of increase in R^*_{it} are: (TI_j) and (K_j); costly, technically complex innovations are most likely to be viewed, initially, with great suspicion, leading potential adopters to require a high risk premium. But, by the same token, as diffusion proceeds, and information improves, there is more scope for a continual revision downwards of the risk premium and revision upwards in R^*_{it}.

Thus let

$$\Psi_j = g_4 \, (\underset{+}{TI_j};\ \underset{?}{CS_j};\ \underset{?}{CC_j};\ \underset{-}{N_j};\ \underset{(+)}{\Pi_j};\ \underset{(+)}{\bar{C}_j};\ \underset{+}{K_j}) \qquad (5.6.4)$$

<u>(v) Ω_j</u>

Ω_j may be defined as a measure of the sensitivity of R_N, H and R* to cyclical variations in the industry concerned. The effect on the profitability of installation (R_N) will presumably be far greater for costly, complex innovations (TI_j, K_j): these, particularly, require intensive usage to function efficiently and, on the other hand, produce the most disruptive influences.[2] The other influence of C_t on $\frac{R^*}{H.R_N}$ stems from the exogeneous pressures it causes, leading to search and revision of goals.

1. Ch. 4, sn. 3(b) and (c).

2. Ch. 3, sn. 8.

Following the arguments used above, the yield of this search and the extent to which entrepreneurs react by changing their investment yardsticks, may be influenced by: CC_j , Π_j , TI_j and N_j.

Thus, let

$$\Omega_j = g_5 (\underset{?}{TI_j}; \underset{?}{K_j}, \underset{-}{CC_j}, \underset{(+)}{\Pi_j}, \underset{-}{N_j}) \qquad (5.6.5.)$$

<u>Summary</u>

The model presented has attempted to incorporate the evidence reported in previous chapters into a simple theory of decision making, based on an underlying behavioural view of the firm. It yields predictions in three areas. First, unlike in past theories of diffusion, the growth curve may follow one of two alternative trends: symmetric or positively skewed S shapes. Further, cyclical fluctuations may be superimposed on these trends. In most past research, curve fitting has been a descriptive device, devoid of much economic interest - a <u>means</u> to the end of generating measures of the speed of diffusion. In this case, however, because the exact shape of the growth curve will depend on the characteristics of the innovation (and perhaps of the industry) involved, curve fitting will be an <u>end</u> in its own right - a test of the model. For instance, one would expect only certain types of innovations (Group A) to diffuse according to a cumulative lognormal trend curve.

Second, whilst the epidemic models simply assume that the speed of diffusion is determined by an ad hoc set of explanatory variables, the above model generates seven parameters, each of which will be seen to influence the speed of diffusion. The descriptions and analysis of the previous chapters suggest a wide range of variables as plausible determinants of five of these parameters.

Third, a cross-section prediction of the model concerns the shape of the Quasi-Engel curve at any point in time, for a given industry. In chapter 7, this curve will be estimated for each of the sample innovations, thus providing a test for two of the major assumptions of the model: namely, that size is an important determinant of firm behaviour and that the error term in (5.2.7) is lognormally distributed.

The following three chapters are concerned with testing these predictions. Chapter 6 fits the alternative growth curves to the diffusion time series for the sample innovations, Chapter 7 estimates the parameters of the Quasi-Engel curves and Chapter 8 provides a cross-industry study of the determinants of the speed of diffusion.

Appendix 1 to chapter 5: The reasons for choosing the linear-in-logs form in section 3.

The choice of linear-in-logs specification for R_{Nit}, H_{it} and R^*_{it} in section 3 is based on two considerations, other than the obvious one of mathematical convenience. First, it allows the effect of S to increase at an increasing or decreasing rate, depending on the value of β, this seems desirable, given many of the arguments used for the inclusion of this variable. Similarly, the multiplicative form ensures that the effect of each of the independent variables (S, θ and ϵ) will be dependent on the levels of the others. Thus, to use just one example, size may only help information-receptiveness, if a number of other characteristics of the firm concerned are favourable e.g. cosmopolite managers. Algebraically, this requires, of course, that $\partial^2 H / \partial S . \partial \epsilon \neq 0$, which will be true for the linear-in-logs form but not, say, for a linear form.

Second, and more specifically, the nature of the scale economies mentioned in chapter 3 suggest this mathematical form.

To use a highly simplified example; consider the case of a firm replacing old technology equipment with the latest vintage of the new innovation. If the new innovation saves only labour[1] and output remains the same before and after adoption, then the payback will be given as:

$$R_{Nit} = \frac{P_{INt} \; K_{Nit}}{W_t \, (L_{Oit} - L_{Nit})} \qquad (5.A1.1)$$

where W_t is the wage rate at time t (assumed roughly constant over the payback period[2]); L_{Nit} and L_{Oit} are the labour inputs per unit of output for firm i, using vintage t new technology equipment and old technology equipment respectively at time t; P_{INt} is the price of one ton of the

1. See Ch. 3, sn. 3.

2. Not too much of an approximation given that many new innovations have paybacks of less than 2 or 3 years (see Appendix one.)

new process of vintage t and K_{Nit} is the capital input (in tons), per unit of output required for vintage t of the new process.

If the innovation has fixed coefficients of production[1]

$$K_{Nit} = a_1 I_i^{b_1} \qquad (5.A1.2)$$

and

$$(L_{0it} - L_{Nit}) = a_{2t} I_i^{b_2} \qquad (5.A1.3)$$

where I_i is the capacity of installation and $b_1 < 0$, $b_2 > 0$ on the basis of the estimates included in section 6 of chapter 3.

Because of learning by doing by the manufacturers (section 7 of chapter 3 and section 6 of chapter 4), for a group B innovation,

$$P_{INt} = P_{IN0}\, e^{g_3 t - g_1 t} \qquad (5.A1.4)$$

where g_3 is the growth rate of wages and g_1 the learning rate

$$\text{and} \quad a_{2t} = a_{20}\, e^{g_2 t} \qquad (5.A1.5.)$$

where g_2 is the learning rate in terms of the specification (or quality) of the innovation.

Substituting in (5.A1.2 - 5.A1.5) into 5.A.1. yields:

$$R_{Nit} = \frac{P_{IN0}\, e^{(g_3 - g_1)t}\, a_1 I_i^{b_1}}{W_t\, a_{20} e^{g_2 t}\, I_i^{b_2}} \qquad (5.A1.6.)$$

and if wages in the consuming industry grow at the same rate g_3 as in the supplying industry,

$$R_{Nit} = \frac{P_{IN0}}{W_0} \frac{a_1}{a_{20}}\, e^{-(g_1 + g_2)t}\, I_i^{b_1 - b_2} \qquad (5.A1.7)$$

Finally, if the capacity of installation is fixed by the size of the firm:

$$I_i = a_3 S_i \qquad (a_3 = 1 \text{ in some cases}) \qquad (5.A1.8)$$

$$R_{Nit} = \frac{P_{IN0} a_1 a_3^{b_1 - b_2}}{W_0\, a_{20}}\, S_i^{(b_1 - b_2)}\, e^{-(g_1 + g_2)t} \qquad (5.A1.9.)$$

1. See ch. 3, sn. 4.

Such an expression is quite compatible, of course, with the form actually chosen for R_{Nit} in equation (5.1.7). The argument is unaltered for Group A innovations, except that now,

$$R_{Nit} = \frac{P_{INO} \ a_1 \ a_3^{\ b_1 - b_2}}{W_0 \ a_{20}} \ S_i^{\ (b_1 - b_2)} \ t^{-(g_1 + g_2)} \qquad (5.A1.10).$$

Appendix 2 to chapter 5: Relaxing some assumptions.

This appendix considers the implications of changing certain assumptions made in the main text; namely that,

$$\frac{d\mu_s}{dt} = \frac{d\sigma_s^2}{dt} = \frac{d\sigma^2}{dt} = 0 \qquad (5.A2.1.)$$

As an example of the effects of changing these assumptions, consider the expression for Q_t for Group B innovations:

$$Q_t = N\,(\Psi t + \Omega C_t \mid -\log\alpha - \beta\mu_s,\ \sigma^2 + \beta^2 \sigma_s^2\,) \qquad (5.5.6b.)$$

The following algebra is simplified by expressing (5.5.6b) in terms of the standard normal distribution and (non-crucially), setting $\Omega = 0$.

$$Q_t = N\left(\frac{\log\alpha + \beta\mu_s}{(\sigma^2 + \beta^2 \sigma_s^2)^{\frac{1}{2}}} + \frac{\Psi t}{(\sigma^2 + \beta^2 \sigma_s^2)^{\frac{1}{2}}} \;\middle|\; 0,\ 1\right) \qquad (5.A2.2.)$$

$$= N\,(q_t \mid 0,\ 1) \qquad \text{where } q_t = a+bt \qquad (5.A2.3.)$$

and

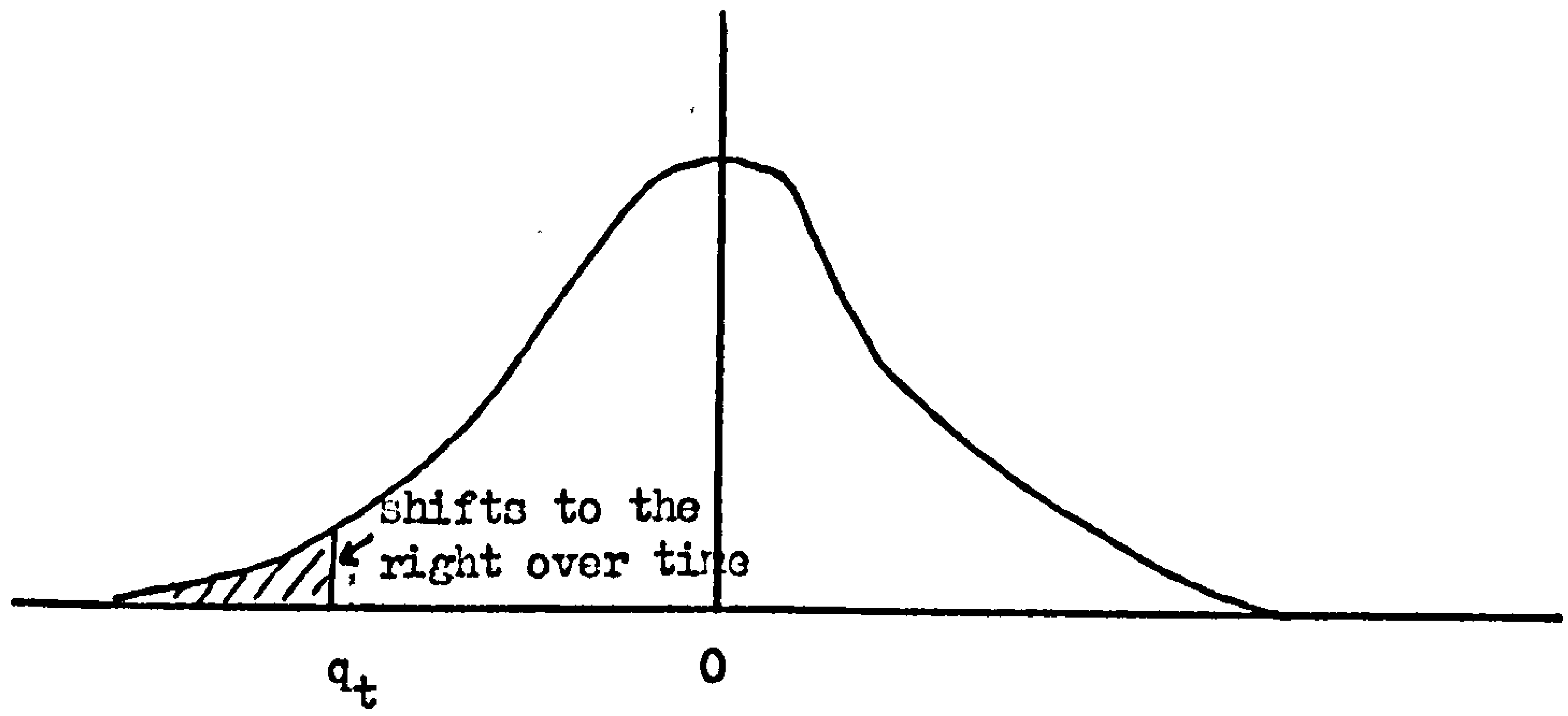
$$a = \frac{\log\alpha + \beta\mu_s}{(\sigma^2 + \beta^2 \sigma_s^2)^{\frac{1}{2}}}, \qquad b = \frac{\Psi}{(\sigma^2 + \beta^2 \sigma_s^2)^{\frac{1}{2}}}$$

So long as assumptions (5.A2.1.) are retained, q_t is a simple linear transform of t and a cumulative normal diffusion curve obtains.

If, however, σ_s^2 or σ^2 increase over time, then both a and b will decrease and the rate of increase q_t will decline over time: $\frac{d^2q}{dt^2} < 0$. If, on the other hand, 6_s^2 or 6^2 decreases over time, then $\frac{d^2q}{dt^2} > 0$.

The implications of these alternative possibilities can be seen intuitively from figure 5.A.2.1.

Figure 5.A.2.1: The implications of variable 6_s^2 and 6^2.



As usual, Q_t is given by the ratio of the shaded area to the total area under the curve; so long as q_t moves to the right at a constant rate over time the diffusion curve will be the cumulative normal. However, if q_t increases at a declining rate, the cumulative curve will be positively skewed with a longer upper tail; if q_t increases at an increasing rate, a negatively skewed cumulative curve will ensue, with a longer lower tail.

As there are, in fact, no a priori grounds for assuming that these variances will change over time,[1] these possibilities are pursued no further.

The third assumption noted in 5.A.2.1. is rather more important. Fortunately, it is also somewhat easier to relax.

Let the geometric mean size of the industry grow at the exponential rate δ[2] Then

$$= e^{\mu_{s0} + \delta t} = e^{\mu_{st}} \qquad (5.A2.4.)$$

1. See footnote 1 to page 20 of this chapter.
2. Given the assumption of no change in 6_s^2 this new assumption implies that the size of the industry as a whole grows at the rate δ.

Substituting into equation (5.5.6a):

$$Q_t = N\left(\Psi \log t + \Omega C_t \,\middle|\, -\log\alpha - \beta\mu_{so} - \delta t\beta\ \ \sigma^2 + \beta^2 \sigma_s^2\right) \quad (5.A2.5.)$$

which might be re-expressed as

$$Q_t = N\left(\Psi \log t + \beta\delta t + \Omega C_t \,\middle|\, - \log\alpha - \beta\mu_{so},\ \sigma^2 + \beta^2\sigma_s^2\right) \quad (5.A2.6.)$$

Similarly for equation (5.5.6b):

$$Q_t = N\left((\Psi + \beta\delta)t + \Omega C_t \,\middle|\, - \log\alpha - \beta\mu_{so},\ \sigma^2 + \beta^2 \sigma_s^2\right) \quad (5.A2.7.)$$

In both cases this presents no problem for estimation purposes. However, for Group A innovations (5.A2.6), the simple lognormal pattern is blurred as 'innovation time' now includes a term in t as well as log t.

Appendix 3 to chapter 5: The expected date of adoption (d_i) as a function of firm size.

Firm i has been assumed to adopt when

$$\frac{R^*_{it}}{(ER_N)_{it}} = 1 \quad (5.A3.1.)$$

That is, when

$$\alpha S_{it}^{\beta}\,\theta_t\,\epsilon_i = 1 \quad (5.A3.2.)$$

Assuming, for simplicity that $\Omega = 0$, the Group A version of (5.A3.2.) is

$$\alpha S_{it}^{\beta}\, t^{\Psi}\, \epsilon_i = 1 \quad (5.A3.3.)$$

and the Group B version is:

$$\alpha S_{it}^{\beta}\, e^{\Psi t} \epsilon_i = 1 \quad (5.A3.4.)$$

If d_i is the value of t which satisfies (5.A3.3) or (5.A3.4) for firm i, then

$$d_i = \alpha^{-1/\Psi}\, S_{it}^{-\beta/\Psi}\, \epsilon_i^{-1/\Psi} \quad \text{(for Group A)} \quad (5.A3.5.)$$

or $$d_i = -(1/\psi)\log\alpha - (\beta/\psi)\log S_i - \frac{1}{\psi}\log\epsilon_i \text{ (for Group B)} \qquad (5.A3.6)$$

However, as was described in section 3(b) of chapter 2, Mansfield and others have estimated relationships of the form

$$d_i = Q\, S_i^{a}\, u_i \qquad (5.A3.7.)$$

where u_i includes a variety of other postulated determinants of d_i. As will be noted, (5.A3.7) is the equivalent of the relationship predicted by this model for Group A innovations i.e. those having a cumulative lognormal diffusion curve. Now, unless S_i or u_i are not lognormally distributed unwittingly, Mansfield has used a mathematical form in this part of his work, which is inconsistent with the symmetrical logistic (close to the cumulative normal) that he uses in his other work on diffusion curve fitting. To have been consistent, he should surely have used the mathematical form:

$$d_i = \log Q + a \log S_i + \log u_i \qquad (5.A3.8)$$

that is, the equivalent of (5.A3.6).

It could be argued, of course, that S_i and u_i need not be lognormally distributed. There is, however, considerable evidence (see Appendix 5) that most size distributions can be at least roughly approximated by the lognormal. Further, u_i represents the effects of six other independent variables which Mansfield includes in the equation in a multiplicative fashion, (see equation 2.3.1 of this thesis), this suggests that a lognormal distribution may be roughly assumed. Indeed, Mansfield makes just this assumption when he omits these other variables from his estimating equation leaving only

$$\log d_i = \log Q + a \log S_i + \log u_i$$

certain statistical problems would emerge if $\log u_i$ were not normally distributed.

Appendix 4. A closer look at the adoption decision.

In the model, the probability of ownership of the innovation (i.e. having adopted) at time t for firm i has been represented by:

$$P\left\{(ER)_{Nit} \leq R^*_{it}\right\} = P_{1t} \text{ (say)} \tag{5.1.2.}$$

An alternative formulation (mentioned on page 5.4.) which has not been pursued is: $P\left\{ER_{Ni\tau} \leq R^*_{i\tau} \quad \text{for some } \tau \leq t\right\} = P_{2t}$ (say) (5.1.2a.)

5.1.2. has been preferred as it provides a far more manageable and fruitful basis for the model. On the other hand, 5.1.2a may be considered preferable, theoretically, in that it avoids the implication that an adopter at $\tau < t$ disadopts (i.e. is no longer an adopter) at a later date if R^* subsequently falls below ER_N.

This implication of 5.1.2. may be considered unreasonable for the following reason. For a firm i to have disadopted by time t, two conditions must be fulfilled in reality: a) an installation of i's using the innovation (as adopted at τ) must at time t no longer use the innovation, b) it must not be profitable for i to re-adopt the innovation at time t. However, the condition for disadoption according to 5.1.2. is merely that $(ER)_{Nit} > R^*_{it}$ (assuming i had adopted at $\tau < t$.) This may be seen as satisfying only condition (b) above. Intuitively, one might suppose that condition (a) would rarely be met when considering successful innovations (i.e. offering substantial cost savings over the old technologies - as all of the sample innovations do.) For (a) to be satisfied, the resale value of the innovation must exceed any disruption costs caused plus future net revenues from operating the innovation.

It seems safe to assume, therefore, that 5.1.2. overestimates the probability of disadoption and underestimates the probability of ownership. (It should be noted that this is a general problem in all economic applications of the probit model to date - see chapter 2, section 6 - as in the present case, past researchers have retained formulations such as 5.1.2. on the implicit assumption that the analytical simplicity they provide outweighs any potential loss in realism.)

Of course, 5.1.2a. also involves a potential approximation as it, in turn, totally excludes the possibility of disadoption; clearly, it assumes 'once an adopter, always an adopter.'

However, at this point it should be noted that no case could be found of any firm actually disadopting any of the sample innovations. Under these circumstances, then, 5.1.2a. is surely preferable and the obvious question which arises is how close an approximation is 5.1.2. to 5.1.2a.?

Defining $P_{3t} = P\left\{(ER)_{Nit} > R^*_{it} \,\middle|\, (ER)_{Ni\tau} \leq R^*_{i\tau} \text{ for some } \tau < t\right\}$ (5.A4.1.)

then $P_{2t} - P_{1t} = P_{3t}.P_{2t}$ (5.A4.2.)

Clearly, a sufficient condition for $P_{3t} = 0$, and thus for P_{1t} to coincide exactly with P_{2t}, is that $(R^*/ER_N)_{it}$ should increase monotonically with t for all i. From (5.2.7.),

$$(R^*/ER_N)_{it} = \alpha S^{\beta}_{it} \; \theta_t \; \epsilon_{it}.$$

Thus, a sufficient condition is that:

$$\frac{1}{\beta}\left(\log\theta_t - \log\theta_{t-1}\right) > -(\log S_{it} - \log S_{it-1}) - (\log\epsilon_{it} - \log\epsilon_{it-1})$$

for all i and t. (5.A4.3.)

(Of course, a number of other, substantially weaker, conditions would suffice, but (5.A4.3.) is probably easier to handle given the available information.)

From the results of chapters 6–8, it is possible to calculate $\frac{1}{\beta}(\log\theta_t - \log\theta_{t-1})$ for all innovations and for all t.

From eqn.(5.5.5.),

$$\frac{1}{\beta}(\log\theta_t - \log\theta_{t-1}) = \frac{\Omega}{\beta}(C_t - C_{t-1}) + \frac{\Psi}{\beta}\log\left(\frac{t}{t-1}\right) \text{ for Group A}$$

$$= \frac{\Omega}{\beta}(C_t - C_{t-1}) + \frac{\Psi}{\beta} \text{ for Group B innovations.}$$

Table 5.A4.1 presents the distribution of minimum and arithmetic mean

values for $1/\beta\ (\log\theta_t - \log\theta_{t-1})$ for the 22 innovations.(Of course, these estimates are based on a model which assumes (5.1.2.) and they may, therefore, be biassed. However, the following analysis is only weakened if the bias, if it exists, is positive - there are no obvious reasons for supposing this to be the case.)

Table 5.A4.1. Observed minimum and arithmetic mean values for $1/\beta\ (\log\theta_t - \log\theta_{t-1})$

Range of values	Minimum	Arithmetic mean
0 - .029	0	0
.03 - .059	3 innovations	0
.06 - .089	2 "	0
.09 - .119	1 "	0
.12 - .199	3 "	2 innovations
.20 - .299	5 "	5 "
.3	8 "	15 "

Sources: Appendix 3, tables 6.6.1.,6.7.2. and 8.3.1.

On the basis of these estimates, then, innovation time (θ_t) increases monotonically with calender time for all 22 innovations. This is a very encouraging result in itself.

Assuming for the moment that ϵ_{it} is constant over time for all i, it can be seen that so long as $\log\ (S_{it-1}/S_{it}) < .03$ for all i and t, then 5.A4.3. will be satisfied for all 22 innovations. This implies that no firm should decline in size by more than 3%. Indeed, so long as no firm declines in size by more than 12% in any year, 5.A4.3. will be satisfied for 16 innovatio (For 16 innovations, $1/\beta(\log\theta_t - \log\theta_{t-1})$ never falls below .12.) Even for the other 6 innovations, a 12% decline in firm size will not contravene 5.A4.3. so long as it occurs at a time when innovation time is proceeding at average or faster pace (column 3 of the table.) In fact for two of the six, $1/\beta\ (\log\theta_t - \log\theta_{t-1})$ only ever falls below .12 in one year in the observation period. Now whilst the 12% figure is, by necessity, arbitrary, it might be fair to assume that the probability of any firm contracting in size by more than this amount will be small: the feasible sets of adopters for the innovations have been defined so as to include only those firms in operation at both the start and the end of the diffusion periods (see Appendix 2.)

Needless to say, the picture becomes more confused, and probably less favourable, if ϵ_{it} is at all variable over time. However, a fairly good case can be made for rough constancy over time. It is perhaps unlikely that the random variables whose influences are represented by ϵ_{it}, (X_{ij} $j=1\ldots r_1$, Y_{ij} $j=1\ldots r_2$, Z_{ij} $j=1\ldots r_3$) will change by much in any firm from year to year. For instance, the technical characteristics of the firm and the educational attainment of its managers should both be fairly fixed over short periods. Moreover, any systematic changes, common to all firms in the industry, will be represented by changing values of θ_t, by definition. Nevertheless, one cannot entirely rule out the possibility of large negative values for $\log\epsilon_{it} - \log\epsilon_{it-1}$ for some i at some t. For most innovations these values would need to be fairly high if 5.A4.3. is not to be satisfied. Unfortunately no data is available on this matter. Having said this, if such values did occur fairly frequently, this would almost certainly result in an upward drift over time in σ^2 (the variance of $\log\epsilon_{it}$.) This would violate one of the crucial assumptions of the model (see eqn.5.3.1.) One of the consequences would then be that the time path of diffusion would not follow the predictions of the model (see Appendix 2 to this chapter.) If this is the case (and the results of the following chapter do not suggest any widespread tendency for the observed time paths to diverge from both the cumulative normal and lognormal curves,) then any approximation involved in the adoption decision equation (5.1.2.) is likely to be far less important than the erroneous assumption of constant σ^2.

Chapter 6 : Time series econometrics.

The time series implications of Chapter 5 are reasonably straightforward: the diffusion growth curve may have a cumulative normal or lognormal trend with cyclical variations superimposed on the trend (see equations 5.5.6. and figure 5.5.5.) In this chapter these predictions are tested against diffusion time series data for the 22 sample innovations, with fairly encouraging results.

Section 1 explains the few remaining steps needed to transform equation (5.5.6) into a form which may be estimated using standard regression techniques. Section 2 suggests, however, that there is a problem of heteroscedasticity which may be satisfactorily overcome by using weighted least squares as an estimating technique. Section 3 explains how the variables have been measured and discusses one or two problems which have emerged. The remainder of the chapter reports the results obtained and assesses the empirical success of the model. As a first step, section 4 evaluates the results of fitting the cumulative normal curve without any cyclical fluctuations and compares these with the logistic curve suggested by Mansfield. Similarly, section 5 evaluates the success of the cumulative lognormal, again omitting cyclical fluctuations. Section 6 reports the results of fitting both curves but also allowing for cyclical fluctuations. Section 7 assesses how far all of these results fit the expectations generated by the theory, and attempts to rationalise the fact that two or three of the sample innovations are not adequately explained by any of the alternative time curves.

1. The estimating equations.

From equation (5.5.6), the probability of any firm, taken at random, having adopted the innovation by time t is given as:

$$Q_t = N(\Psi \log t + \Omega C_t \mid -\log\alpha - \beta\mu_s,\ \sigma^2 + \beta^2\sigma_s^2) \quad (5.5.6a)$$

or

$$Q_t = N(\Psi t + \Omega C_t \mid -\log\alpha - \beta\mu_s,\ \sigma^2 + \beta^2\sigma_s^2) \quad (5.5.6b)$$

Neither of these forms may be estimated directly, but a simple transformation, using the standard normal distribution, N (0,1), easily overcomes the problem.

(5.5.6) may be re-written, quite generally, as:

$$Q_t = N\left\{ \frac{\log\alpha + \beta\mu_s + \Psi \log t + \Omega C_t}{(\sigma^2 + \beta^2\sigma_s^2)^{\frac{1}{2}}} \middle| \; 0, 1 \right\} \qquad (6.1.1a)$$

$$\text{or } Q_t = N\left\{ \frac{\log\alpha + \beta\mu_s + \Psi t + \Omega C_t}{(\sigma^2 + \beta^2\sigma_s^2)^{\frac{1}{2}}} \middle| \; 0, 1 \right\} \qquad (6.1.1b)$$

That is: $Q_t = N\,(\,a + b\log t + c\,C_t | 0,1)$ (6.1.2a)

or $Q_t = N\,(\,a + bt + cC_t\,| 0,1)$ (6.1.2b)

$$\text{where} \quad a = \frac{\log\alpha + \beta\mu_s}{(\sigma^2 + \beta^2\sigma_s^2)^{\frac{1}{2}}}; \quad b = \frac{\Psi}{(\sigma^2 + \beta^2\sigma_s^2)^{\frac{1}{2}}}; \quad c = \frac{\Omega}{(\sigma^2 + \beta^2\sigma_s^2)^{\frac{1}{2}}} \qquad (6.1.3)$$

Now let $z_t = a + b\log t + cC_t$ (6.1.4a)

or $z_t = a + bt + cC_t$ (6.1.4b)

In which case, of course, z_t is the standard <u>normal equivalent deviate</u>[1] of Q_t :

$$Q_t = N\,(z_t | 0,1) = \int_{-\infty}^{z_t} (2\pi)^{-\frac{1}{2}}\; e^{-u^2/2}\; du. \qquad (6.1.5)$$

Equations (6.1.4a) and (6.1.4b) form the basis of the time series estimation of this chapter. z_t may be read off from normal distribution tables, given known values of Q_t, and thus, this transformation permits the use of simple linear estimation techniques.

<u>2. The method of estimation.</u>

At first sight there seems no reason for not using ordinary least squares to estimate equations 6.1.4. However, it must be remembered that the theory has been couched in terms of probabilities. Whilst it might be quite acceptable, for empirical purposes, to equate the probability of a randomly selected firm having adopted (Q_t) with the <u>actual</u> proportion of firms having

1. As defined by J.H. Gaddum (1953), see Aitchison and Brown, op.cit., p.68.

adopted ($\bar{Q}_t$), this does entail a non-trivial assumption.

Following convention[1], $\bar{Q}_t$ is assumed to be a random variable binomially distributed around a mean Q_t with variance of $Q_t(1-Q_t)/n_t$, where n_t is the sample size at time t.

Quite clearly, however, this violates one of the assumptions of the normal linear regression model, namely that of homoscedasticity (constant variance of the disturbance term.) Thus, it would be incorrect to estimate equations (6.1.4) using ordinary least squares.

In fact this problem is always encountered when using probit analysis. In the past, two alternative estimators have been advocated and used: maximum likelihood [2] and the so-called 'minimum normit χ^2.' [3] Both methods have been shown to yield asymptotically efficient estimates,[4] but their small sample properties are less unequivocal. Berkson,[5] using a hypothetical small sample experiment, found that both methods yielded bias which was quite small but that the variances of the estimated coefficients were smaller for the normit χ^2. However, Finney [6] also reports some hypothetical experiments of Cramer's which came out in favour of maximum likelihood.

Both methods are equivalent, then, at the limit but one or the other may perform better for finite samples. As there are no definite grounds for preferring one to the other, the minimum normit χ^2 has been used for all time series regressions in this study, as it requires slightly less computation.[7]

1. see, for instance, Bonus (op.cit.) pp. 659 - 660.

2. see, for instance, D.H. Finney, 'Probit analysis, a Statistical Treatment of the Sigmoid Response Curve.' Cambridge: The University Press (1947) and Aitchison and Brown (op.cit.) Chapter 7.

3. J. Berkson, 'Estimate of the integrated normal curve by minimum normit chi-square.' Journal of the American Statistical Assoc., June 1955, p.529.

4. W.F. Taylor, 'Distance Functions and regular best asymptotically normal estimates.' Annals of Mathematic Statistics, 24 (1953) p.85 - 92.

5. J. Berkson, 'Tables for use in estimating the normal distribution function by normit analysis.' Biometrica, Vol.44, 1957.

6. Finney, (op.cit.)

7. Mansfield (1968), op.cit., p.141 also prefers the minimum logit method (the equivalent to minimum normit χ^2 when estimating logistic curves.)

The basis of the method lies in minimising the quantity:

$$\sum n_t \left(W_t^2 / \bar{Q}_t(1-\bar{Q}_t) \right)(\bar{z}_t - \hat{z}_t)^2 \qquad (6.2.1)$$

where $\bar{z}_t$ is the normal equivalent deviate (or normit) of $\bar{Q}_t$ (observed diffusion) and $\hat{z}_t$ is the predicted normal equivalent deviate.

W_t is the ordinate of the standard normal curve at the point where its area is divided into $\bar{Q}_t$ and $(1-\bar{Q}_t)$.

Minimising (6.2.1) can be effected by using weighted least squares regressions with each observation being weighted by $\dfrac{(n_t W_t^2)}{\bar{Q}_t(1-\bar{Q}_t)}$.[1]

Thus minimum normit χ^2 is merely a special case of the standard method of correcting for heteroscedasticity: namely to weight the observations by the inverse of the variance of the theoretical error term.[2]

Strictly speaking, of course, the quantity which should be minimised is

$$\sum \frac{n_t}{Q_t(1-Q_t)} \cdot (\hat{Q}_t - \bar{Q}_t)^2 \qquad (6.2.2)$$

But Q_t is unknown, of course (and also the data has been transformed into normits.) A formal argument of the case for using (6.2.1) as a first order approximation to (6.2.2) is given by both Berkson[3] and Finney[4]. On a more intuitive level, however, the argument can be summarised fairly simply. $\dfrac{n_t}{\bar{Q}_t(1-\bar{Q}_t)}$ is an estimate of the accuracy of the approximation of the probability by the observed frequency and W_t^2 may be thought of as a measure of the accuracy of the transformation from proportions to normits. As can be seen from the diagram, for small changes in $\bar{Q}_t$ at low and high levels, the change in $\bar{z}_t$ is far greater proportionately. Consequently, an error in

1. Berkson, (1957), op.cit., provides tables of values for these weights.

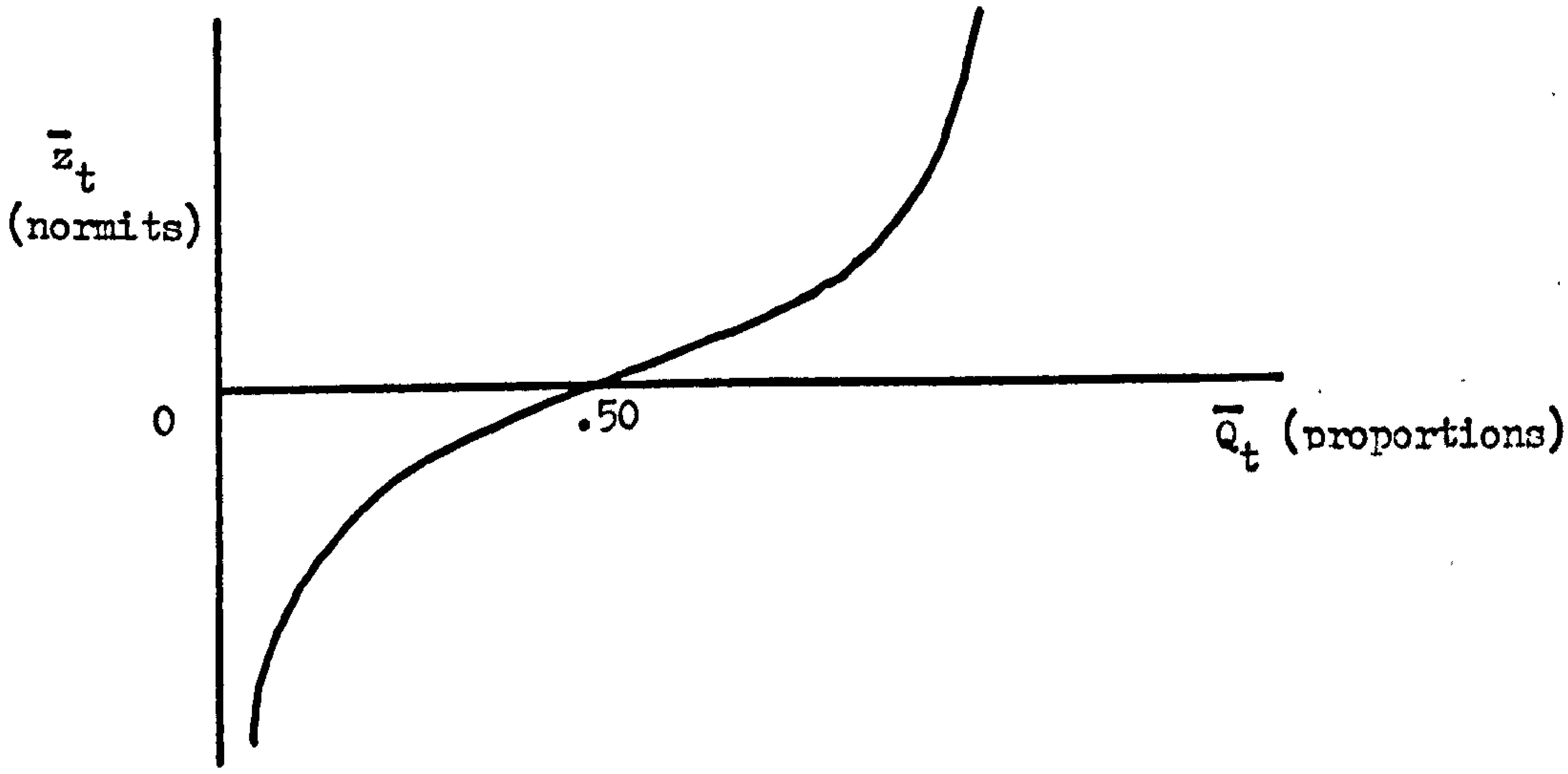
2. See, for instance, J. Johnston, 'Econometric Methods,' 1st edition, McGraw Hill, New York (1963), pp.207-211.

3. Berkson, op.cit.

4. Finney, op.cit.

$\bar{Q}_t$ at low or high levels would result in a far larger error in $\bar{z}_t$ than would be the case for middling levels of $\bar{Q}_t$.

Figure 6.2.1: Normits and proportions.



Given that $\bar{Q}_t = \int_{-\infty}^{\bar{z}_t} (2\pi)^{-\frac{1}{2}} . e^{-u^2/2} du.$,

$$\frac{d\bar{z}_t}{d\bar{Q}_t} = (2\pi)^{\frac{1}{2}} e^{-\bar{z}_t^2/2} = \frac{1}{W_t} .$$

Therefore, the expected variance in $\bar{z}_t$ caused by a small deviation in $\bar{Q}_t$ is given by $^1/W_t^2$ and the weight for the observed error $(\bar{z}_t - \hat{\bar{z}}_t)$ is given by the inverse of $\frac{\bar{Q}_t(1-\bar{Q}_t)}{n_t} . \frac{1}{W_t{}^2}$.

Berkson suggests, as a measure of goodness of fit, the value of (6.2.1), which can be tested for significance using χ^2 tables. However, this is a rather weak test and inappropriate for evaluating alternative time curves. Of the commonly used test statistics, the Durbin-Watson is perhaps the best suited to testing for appropriate specifications of mathematical form, although a satisfactory $\bar{R}^2$ or χ^2 statistic is also necessary for acceptance.

3. Measurement of variables.

(i) The dependent variable: Appendix 2 presents the data from which $\bar{z}_t$ has been computed. $\bar{Q}_t$ measures the observed cumulative number of adopters as a proportion of the total number of potential or feasible adopters[1] in the industry at yearly intervals. $\bar{Q}_t$ is transformed into $\bar{z}_t$ by using standard normal distribution tables.

(ii) The cyclical variable, C_t : as mentioned in the previous chapter the empirical measure of C_t used is capacity usage in the industry as a whole in year t. This, in turn, has been measured as the ratio of actual output to total capacity. In general, of course, data on capacity is rarely available and so capacity figures have been computed using the most simple Wharton school technique for every industry in this sample. It is assumed that full capacity working occurs at the peak of each cycle and estimates of capacity for the intervening years are derived by using straight line interpolations between the peaks. Thus capacity working may vary between 0 and 1. However, for these purposes C_t has been defined as the deviation of capacity working in year t from the level in year 0 (i.e. $C_o = 0$.) This will not alter the estimates $\hat{c}$ (as var(C) is unaffected), and is merely a way of correcting for differences between industries in the typical level of C which would otherwise be inversely related in $\hat{a}$.

Two deficiencies of this variable both stem from poor quality data. First, better methods of computing capacity from production figures are available but require ancillary knowledge on capital stock and/or fuel consumption; unfortunately, such data is unavailable at the level of industry disaggregation necessary in this study. Second, for some industries, even production data was unavailable and sales or demand figures had to be used in their stead.[2]

Three alternative measures of C_t were used in place of capacity usage:

1. The feasible set for each innovation was computed on purely technical grounds, see Appendix one. $\bar{Q}_t$ is the equivalent, then, of Mansfield's $(m/n)_t$, see Chapter 2, sections 1 and 2.

2. See Appendix 3 for more detail.

industry unemployment, gross deflated investment and demand. But, as reported in appendix three, these were based on very poor quality and inappropriate data and the results obtained using them in the estimating equations are not reported. For the record, none of these three variables proved to be significant for more than two or three innovations and in each case a more satisfactory fit was obtained using capacity usage.

(iii) Time, t: when estimating the cumulative normal form, the date of the first adoption is designated as t = 1. In this case, the choice of origin does not influence $\hat{b}$ or $\hat{c}$,[1] (in eqn.6.1.4b.) but $\hat{a}$ will be affected as the least squares formula is, of course:

$$\hat{a} = \bar{z} - \widehat{b\bar{t}} - \widehat{c\bar{C}} .$$

Thus, had t been measured by calendar years such that $\bar{t}$ was, say, 1960, $\hat{a}$ would have been far smaller. Given that the choice of the time origin is arbitrary, then so too will be the absolute level of $\hat{a}$ for any innovation. However, so long as the same convention is used for all innovations, the variance of $\hat{a}$ will be insensitive to how time is measured and the cross-section analysis of $\hat{a}$ in the next chapter is still valid.

A more serious problem is encountered, however, when estimating the cumulative lognormal form. In this case, the variance of log t <u>will</u> depend on the choice of origin and, therefore, so will $\hat{b}$. To put the point another way, as the cumulative lognormal (unlike the cumulative normal) is constrained, by its mathematical form, to pass through the origin, this rules out certain choices for 'the start of time' as unreasonable. Consider the following hypothetical example of an innovation which is adopted by the first firm in 1960, followed by 2 firms in 1961, 4 more in 1962 and so on. If there are 100 firms in the industry and time is measured by calendar years, the curve will be fitted to the following points: (0;0), (1960;1), (1963;3), (1962;7) and so on. Obviously both $\hat{b}$ and the overall fit would be different from those obtained measuring time as for the cumulative normal and generating the points: (0;0), (1;1), (2;3), (3;7). Although these differences will be

1. Nor the overall goodness of fit.

less pronounced for other choices of origin (e.g. assuming first adoption takes place in year 2), they may still be important. Of course, were the choice of origin one year before first adoption based on the definite knowledge that the innovation was, indeed, first available at that date, then the problem would disappear. However, in general, it is impossible to tell precisely when any given innovation was first made commercially available.

Thus, since the choice of origin must be arbitrary, a number of alternative origins were used for each innovation. Limits were imposed, however, on the lag implied between availability and first adoption: it was fairly obvious from the technical literature that in no case was any innovation available for more than four years before it was first adopted. Therefore, for each innovation, first adoption is assumed to have taken place at any one of the following: t = 5;4;3;2;1;.7;.5;.1 or .01. As it happens, for half the sample, the fit is poor regardless of 'when time started' but for others, notable differences (particularly in the Durbin-Watson statistic) do occur when using different origins. In these cases, the choice of origin which provides the best overall fit is reported.[1]

4. Results - the simple cumulative normal.

As a first step, the line

$$\bar{z}_t = a + bt \qquad (6.4.1)$$

is fitted to the diffusion data for all 22 innovations. The twin assumptions made initially, then, are of cumulative normal growth paths without cyclical variations, that is, that all innovations are Group B with Ω (and thus c) = 0.

1. In this instance, because the choice of origin is arbitrary anyway, this procedure seems quite valid. There is no question of it leading to the acceptance of the lognormal hypothesis when, in fact, the data follows a symmetrical or negatively skewed S-shape. In such cases, the fit will always be better for the cumulative normal, no matter how time is measured in the lognormal; unless $\hat{b} < 0$ (which is impossible, given that diffusion increases monotonically with time), (6.1.4a) always implies a positive skew.

The graphical counterpart of this line, in terms of $\bar{Q}_t$ is given by figure 5.5.5., part (d).

As a matter of interest, the linear transformation of the logistic curve[1] was also fitted: (using weighted least squares.)

$$\log\left(\frac{\bar{Q}_t}{1-\bar{Q}_t}\right) = a' + b't . \qquad (6.4.2)$$

This provides a useful yardstick against which to judge the results of this chapter, as well as a further test of Mansfield's epidemic model. Given the close similarity between the shapes of the logistic and the cumulative normal,[2] it would be surprising if any marked differences in 'goodness of fit' emerged between the two alternative hypotheses.

Table 6.4.1. presents the estimated coefficients, $\bar{R}^2$ and Durbin-Watson (D-W) statistics for all 22 innovations.

Judged by corrected R^2 and t statistics, both logistic and cumulative normal offer fairly convincing explanations. For the cumulative normal, in 19 cases $\bar{R}^2 > .75$ and for 17 of the 22 innovations, the $\bar{R}^2$ for the logistic exceeds .75. All 88 coefficients are significantly different from zero using a 5% test and the vast majority are also significantly different at the 1% level.[3] This is hardly surprising, however, and is certainly not conclusive. As all of the sample innovations are successful, diffusion must rise monotonically with time, and given that $\bar{Q}_t$ is constrained to lie between 0 and 1, an upward sloping S shape is almost assured. Therefore, whilst high $\bar{R}^2$ and t statistics are necessary conditions for non-rejection of hypotheses, they are certainly not sufficient.

In this particular case, the D-W statistic which tests for non-randomness of residuals is a more appropriate test. If, for instance, the true relationship is a skewed S shape, then fitting a symmetrical S shape to the data should generate a set of residuals positively serially correlated (and thus a D-W statistic below the lower critical level.)

1. see equation 2.1.4, Ch.2. sn.1.
2. See N. Johnson and S. Kotz, "Continuous univariate distributions -2", Houghton Mifflin, Boston(1970) pp.1-18.The main difference is the relatively longer tails of the logistic.
3. The estimated intercept terms for the logistic, $\hat{a}$, are not shown in the table as they play no part in the ensuing cross-section analysis(following Mansfield)

Of course, the D-W test is also not sufficient on its own: it would hardly be satisfactory for the residuals to be randomly distributed over time if at the same time a large proportion of the variance in $\bar{Q}_t$ is not 'explained' by the equation.

Thus a fair yardstick by which to judge each result is the requirement that it should exhibit a high $\bar{R}^2$, significant coefficients and the absence of positive serial correlation.

Table 6.4.1. Results using the simple cumulative normal curve.

Innovation	Cumulative Normal $\hat{b}$	$\hat{a}$	$\bar{R}^2$	D-W	Logistic $\hat{b}'$	$\bar{R}^2$	D-W
A) Cumulative normal acceptable set.							
T.K.	.093 (29.07)	-2.274 (52.18)	.978	1.647	.176 (20.76)	.958	1.189 (*)
V.M.	.190 (10.41)	-1.997 (14.39)	.906	1.410	.330 (10.09)	.901	1.434
W.S.B.	.076 (11.54)	-2.083 (32.45)	.904	1.296 (*)	.156 (11.51)	.903	1.283 (*)
NC.TURB.	.161 (11.72)	-1.603 (10.01)	.906	1.147 (*)	.281 (11.56)	.904	1.223 (*)
A.T.L.	.179 (12.67)	-1.174 (10.66)	.924	1.786	.292 (12.00)	.916	1.777
P.C.B.C.	.140 (5.35)	-2.399 (20.36)	.816	1.847	.359 (4.65)	.765	1.474
B.O.P.	.151 (7.11)	-1.549 (10.85)	.836	1.403	.250 (6.36)	.777	1.223 (*)
T.C.	.046 (6.00)	-1.498 (21.30)	.708	1.261 (*)	.083 (6.02)	.710	1.184 (*)
G.A.	.204 (7.65)	-1.349 (8.45)	.863	1.146 (*)	.320 (5.31)	.849	1.014 (*)
C.T.	.186 (6.52)	-2.137 (11.32)	.818	1.128 (*)	.320 (7.26)	.743	.776 *
F.	.307 (9.41)	-2.129 (10.29)	.906	1.119 (*)	.506 (7.74)	.866	.793 *
S.F.	.213 (9.04)	-1.664 (11.44)	.899	1.022 (*)	.346 (7.87)	.870	.885 *
B) Cumulative normal unacceptable set							
NC.PP.	.131 (7.26)	-1.297 (11.99)	.849	.910 *	.213 (6.57)	.820	.802 *
C.C.	.108 (4.58)	-1.596 (9.02)	.633	.576 *	.174 (3.86)	.540	.471 *
S.P.	.202 (6.32)	-2.249 (10.49)	.851	.548 *	.345 (4.93)	.776	.334 *
A.S.B.	.056 (7.79)	-2.020 (28.04)	.796	.375 *	.108 (7.39)	.778	.346 *
A.D.H.	.044 (3.74)	-1.371 (11.08)	.429	.272 *	.068 (3.53)	.395	.272 *
E.H.	.057 (8.39)	-1.684 (17.64)	.765	.212 *	.095 (7.38)	.714	.231 *
S.L.	.116 (16.56)	-2.910 (33.67)	.944	.334 *	.260 (17.12)	.948	.335 *
NC.TURN.	.151 (9.92)	-2.282 (13.81)	.865	.475 *	.269 (9.01)	.841	.442 *
S.P.C.	.113 (10.04)	-2.258 (18.52)	.868	.681 *	.217 (9.25)	.848	.677 *
V.D.	.150 (14.53)	-2.157 (21.17)	.937	.953 *	.272 (14.47)	.937	1.241 (*)

Notes to Table 6.4.1.

1.Bracketed figures underneath coefficients denote t statistics.

2.* underneath D-W statistics denotes significant positive serial correlation of the disturbance term at the 5% level.

3.(*) underneath D-W statistics denotes that the test has proved inconclusive at the 5% level.

4.For definitions of the two sets, see the text.

On this basis, the cumulative normal performs satisfactorily for 12 of the 22 innovations, in that $\bar{R}^2$ exceeds .7 and the D-W statistic does not confirm the existence of significant serial positive correlation of the error term at the 5% level.[1] In the table, these 12 are designated as the 'cumulative normal acceptable set'. For the other 10 innovations, whilst t values and $\bar{R}^2$ are still very high, the D-W values suggest strongly that autocorrelation exists. An obvious explanation for this second set is that the mathematical form has been mis-specified and/or a significant variable has been omitted from the equation. The following sections attempt to answer these possibilities by fitting the cumulative lognormal and including the cyclical variable C_t, in the equation.

Rather surprisingly, the logistic curve offers an inferior explanation for virtually all of the innovations. For 19 of the 22, $\bar{R}^2$ and the t statistic for $\hat{b}$ are higher for the cumulative normal than for the logistic, but probably not significantly so in most cases. More interestingly, for only 3 of the innovations can significant autocorrelation be ruled out and for 12 (as opposed to 10 for the cumulative normal) there is conclusive evidence of significant autocorrelation.

At the very least, then, the cumulative normal definitely offers no worse overall explanation than the logistic; as can be seen from the table, only 8 of the cumulative normal acceptable set could be defined as logistic acceptable.

5. Results : The simple cumulative lognormal.

In table 6.5.1. the results are reported of fitting the line

$$\bar{z}_t = a + b \log t . \qquad (6.5.1)$$

Thus again the cyclical influences are omitted, but this time a cumulative lognormal growth curve is hypothesised. The graphical counterpart of this equation is figure 5.5.5.,(part (b)).

1. Although in 7 of the 12 the D-W test is inconclusive: that is, it is not certain that autocorrelation does not exist.

As in table 6.4.1. the usual statistics are reported. In addition, for each innovation, column 5 reports the value given to t at the time of the first adoption. As explained in section 3 this is decided on the basis of the choice of origin which provides the best fit. Thus for TC, the time series for t reads: .3; 1.3; 2.3; 3.3; and so on: this implies that the innovation was first adopted 4 months after it was first available.[1] Having said this, for exactly half of the innovations (usually those for which the lognormal is obviously inappropriate), differences in the choice of origin have virtually no effect on goodness of fit; in those cases, arbitrarily, the results achieved assuming $T_1 = 1$ are reported.

As before, $\bar{R}^2$ and t statistics are nearly always high; 18 of the 22 $\bar{R}^2$ exceed .75 and all coefficients are significantly different from zero even at the 1% level. But also as before, the D-W statistics often suggest significant autocorrelation. 13 of the 22 fall into a <u>cumulative lognormal acceptable set</u> with high $\bar{R}^2$, t statistics and no conclusive evidence of positive serial correlation in the error term.[2] Of the other nine innovations, eight quite clearly do not have lognormal diffusion curves as both $\bar{R}^2$ and D-W statistics are relatively low. In the case of EH, however, whilst the D-W value is below the lower limit, the $\bar{R}^2$ is very high, especially when compared with the cumulative normal and logistic fits shown in table 6.4.1.

A comparison of the cumulative lognormal and logistic over all 22 innovations again suggests that the logistic comes off second best. For 13 of the 22

1. It will not be argued, however, that the variance between innovations in the 'optimal origin' has any definite economic meaning. Thus the (tenuous) implication that SP, for instance, was available for 3 years before any firm adopted will not be pursued.

2. Although for 2 of these innovations the test is inconclusive and for one other (NCPP) the D-W suggests negative autocorrelation. It is tempting to dismiss the latter as a statistical fluke. Given that the data series is not in first difference terms, it seems unlikely that this result is due to mis-specification of form. Indeed no obvious explanation suggests itself.

Table 6.5.1. Result using the simple cumulative lognormal curve.

Innovation.

a) Cumulative lognormal acceptable set	$\hat{b}$	$\hat{a}$	$\bar{R}^2$	D-W	T_1=
TC.	.238 (8.01)	-1.546 (26.36)	.816	1.467	.3
GA.	.766 (14.65)	-1.334 (16.08)	.959	2.165	.5
CT.	1.328 (8.75)	-3.719 (8.75)	.892	1.695	3.0
F.	2.107 (11.06)	-4.571 (11.70)	.930	1.505	3.0
SF.	1.231 (15.17)	-2.701 (17.73)	.962	1.738	2.0
ASB.	.339 (29.47)	-2.168 (89.06)	.983	2.089	.3
ADH.	.311 (8.56)	-1.570 (20.13)	.816	2.150	.01
NCPP.	.509 (20.38)	- 1.334 (32.93)	.979	3.358 **	.7
CC.	.742 (6.63)	-2.323 (10.17)	.792	1.068 (*)	2.0
SP.	1.494 (9.38)	-4.080 (12.14)	.926	1.663	3.0
ATL.	.882 (8.54)	-1.479 (7.58)	.845	1.251 (*)	1.0
PCBC.	.424 (4.88)	-2.375 (19.09)	.784	1.830	1.0
BOP.	.752 (7.61)	-1.859 (10.04)	.815	1.358	1.0
b) Cumulative lognormal unacceptable set					
EH.	.549 (26.40)	-2.250 (43.97)	.971	.877 *	.3
TK.	.772 (12.94)	-2.963 (19.70)	.897	.601 *	1.0
VM.	.877 (5.99)	-2.260 (8.06)	.755	.987 *	1.0
WSB.	.441 (7.32)	-2.302 (17.92)	.786	.945 *	1.0
NCTB.	.890 (6.53)	-1.945 (7.02)	.744	.630 *	1.0
SL.	.725 (6.44)	-3.265 (11.95)	.712	.285 *	1.0
NCTURN.	.866 (5.90)	-2.641 (7.94)	.686	.356 *	1.0
SPC.	.647 (5.95)	-2.525 (10.26)	.690	.426 *	1.0
VD.	.854 (6.96)	-2.560 (9.64)	.768	.369 *	1.0

Notes to Table 6.5.1.

1. Bracketed figures underneath coefficients denote t statistics.
2. * and (*) as for table 6.4.1.
3. For definitions of the two sets see the text.
4. T_1 indicates the value given to t at the time of the first adoption. Thus the first observation on t for TC is .3, implying that the innovation was available for .3 of a year before it was first adopted. These values of T_1 have been chosen using the criterion outlined in the text.
5. ** denotes significant negative serial correlation of the disturbance term at the 5% level.

innovations, $\bar{R}^2$ and the t statistic for $\hat{b}$ are higher for the cumulative lognormal than for the logistic, and for only 9 is there conclusive evidence of significant positive autocorrelation as opposed to 12 for the logistic.

A comparison of the cumulative lognormal and normal curves is not attempted as the model suggests that each will be applicable for different types of innovation (Group A and Group B as defined earlier.) A brief glance at the two tables will confirm, however, that each curve performs better for roughly the same number of innovations.

6. Results : the inclusion of cyclical variations.

Table 6.6.1 reports the results achieved when both 6.4.1 and 6.5.1 are extended to include the variable C_t:

$$\bar{z}_t = a + bt + cC_t \qquad (6.6.1)$$

and $$\bar{z}_t = a + b\log t + cC_t \qquad (6.6.2)$$

That is, the cumulative normal and lognormal curves with cyclical fluctuations superimposed. The graphical counterparts of these equations are shown in figure 5.5.5, parts (c) and (a).

Before considering this table, one added complication must be mentioned. The model is based on the decision to adopt yet data is available, in practice, only for the date at which firms first started to use the new innovations. The extent of the time lag between the decision to adopt and the first use will obviously vary between innovations, (depending on how long they take to install and whether any experimentation period is necessary.) Thus, as C_t refers to the state of the trade cycle at the time of decision, in these regressions it should be lagged. In general, there is insufficient information available to decide, ex-ante, which lag is most appropriate for each innovation, therefore, three alternatives have been tried in each case: C_{t-1}, C_{t-2} or C_t (where the time units are measured in years.)

In general, the inclusion of this variable produces only minor improvements in the fits reported in tables 6.4.1 and 6.5.1. In the majority of regressions the coefficient, $\hat{c}$, is insignificantly different from zero. Table 6.6.1

reports only those cases in which the explanatory power of the original equation (as measured by $\bar{R}^2$) is improved by the inclusion of the extra variable.[1] Even for these results, however, C_t is insignificant more often than not.

More specifically, C_t has a significant influence (at the 5% level) in equation (6.6.1) for five of the 12 innovations in the 'acceptable cumulative normal set'; using the weaker 10% t test, another innovation (VM) may be added to the five. For only two of the thirteen innovations in the 'acceptable cumulative lognormal set' does C_t have a significant influence in equation (6.6.2); again another innovation (SP) may be added if the weaker 10% test is used.

For the innovations outside of these two sets, C_t never has a significant influence, even using a 10% test. Interestingly, however, for two innovations (NCPP) and (VD) there is no longer conclusive evidence of positive autocorrelation when fitting the cumulative normal curve.

Probably the most notable positive aspect of this table is that the 'best' lag for each innovation is generally in line with expectations of the probable installation periods. BOP, TK, VM, VD, PCBC, and, to a lesser extent, CT are all large pieces of equipment which would require a relatively long time to build, install and balance with adjacent plant, and the two year lag suggested by the regressions seems quite reasonable. NCPP, SPC and TC are less fundamental and 'lumpy', but each probably requires non-trivial installation and testing periods of about one year. F, SP, EH and ASB, on the other hand, are cheap, mainly supplementary innovations which can probably be installed almost instantaneously. Only GA and SF (relatively minor innovations for which no lag might be expected) run against expectations, with a one year lag providing the best fit.

1. This is roughly equivalent to reporting only those regressions for which the computed t statistic on $\hat{c}$ exceeds 1.

Table 6.6.1. Results using cumulative normal and lognormal diffusion curves with cyclical fluctuations.

Innovation	Best lag i on C_{t-i}	cumulative normal $\hat{b}$	$\hat{a}$	$\hat{c}$	$\bar{R}^2$	D-W	cumulative lognormal $\hat{b}$	$\hat{a}$	$\hat{c}$	$\bar{R}^2$	D-W
PCBC	2	.142	-2.377	-2.967	.951	1.551					
		(10.49)	(19.58)	(3.46)							
BOP	2	.148	-1.511	-1.662	.875	1.946	.782	-1.858	-0.990^{+}	.841	1.607
		(8.47)	(10.02)	(2.30)			(8.23)	(8.47)	(1.33)		
CT	2	.205	-2.395	7.178	.892	1.982	1.359	-3.770	3.336^{+}	.902	1.910
		(7.86)	(8.10)	(2.41)			(8.31)	(6.68)	(1.31)		
GA	1	.192	-1.366	1.267	.940	1.673	.753	-1.195	4.376^{++}	.962	1.439
		(8.43)	(6.53)	(3.19)			(10.70)	(15.59)	(1.21)		
F	0	.310	-2.018	-5.688	.943	1.302	2.152	-4.626	-5.489	.968	2.368
		(12.15)	(9.42)	(2.36)			(16.44)	(10.30)	(3.06)		
SF	1	.220	-1.618	-2.277^{++}	.899	.757	1.268	-2.691	-3.046	.986	1.904
		(8.96)	(10.23)	(1.00)		*	(24.50)	(16.19)	(3.53)		
TK	2	.096	-2.279	.400^{++}	.978	1.747					
		(25.81)	(37.46)	(1.14)							
VM	2	.187	-1.900	.932^{+}	.922	1.652					
		(11.20)	(14.57)	(1.69)							
TC	1						.238	-1.572	.608^{+}	.828	1.321
							(8.28)	(24.96)	(1.36)		(*)
SP	0	.208	-2.199	-3.439^{+}	.852	.624	1.487	-4.138	-3.069^{+}	.951	1.611
		(6.52)	(10.38)	(1.23)		*	(10.55)	(12.01)	(1.73)		
NCPP	1	.144	-1.166	-2.527^{+}	.865	1.201					
		(7.32)	(10.62)	(1.35)		(*)					
ASB	0						.337	-2.153	.158^{+}	.984	1.776
							(29.73)	(69.01)	(1.21)		
EH	0						.550	-3.253	.257^{+}	.971	.946
							(26.78)	(22.12)	(1.23)		*
SPC	1	.115	-2.251	-1.466^{+}	.874	.703					
		(10.34)	(16.81)	(1.27)		*					
VD	2	.153	-2.130	.739^{+}	.940	1.225					
		(14.75)	(15.05)	(1.27)		(*)					

Notes:
1. t statistics in brackets.
2. * and (*) as for table 6.4.1.
3. + denotes a coefficient significantly different from zero at the 25% level but not the 5% level.

   ++ denotes a coefficient not significantly different from zero at the 25% level.

The most important conclusion of this section must be, however, that the trade cycle does influence the diffusion growth curve but only in a small minority of cases. For four innovations (PCBC, BOP, F and SF), periods of high economic activity tended to slow down diffusion, whilst low economic activity increased its speed. But for two other innovations (GA and CT) the opposite effect obtained. For the remaining sixteen innovations, however, there is no significant evidence either way. Given the generally poor quality of data used to measure C_t, however, it would be wrong to generalise these results too strongly.

Finally, a tail piece conclusion which may be drawn is that the estimates of a and b are fairly insensitive as to whether C_t is included or excluded from the equations. This suggests that C_t is not seriously correlated with either time or the constant term and therefore, $\hat{a}$ and $\hat{b}$ should not be biassed in tables 6.4.1 and 6.5.1 by the exclusion of C_t at least.

7. The time series success of the model.

As the model developed in chapter 5 generated four alternative hypotheses as to the shape of the diffusion growth curve, its explanatory success must be judged in the light of all the results reported in the previous three sections. Thus, the fact that the simple cumulative normal curve, for instance, proved to be an acceptable explanation for only just over half of the innovations is only disappointing if none of the other alternatives provided an acceptable explanation for the other half.

In comparison with the simple logistic curve, the model clearly provides a better overall explanation. Even discounting the regressions in which C_t was included, either the simple cumulative or lognormal curves provided better fits (as measured by $\bar{R}^2$ and t statistics) than the logistic for twenty one of the twenty two innovations.[1] Similarly, whilst the logistic produced

1. This success rate for the cumulative normal and lognormal (i.e. $^{21}/_{22}$) is significantly different from $^2/_3$ at the 95% level. That is the hypothesis that this result could have come from a binomial distribution with mean 2/3 may be rejected at the 95% level. 2/3 might have been expected if each curve was equally applicable.

significant positive autocorrelation in twelve cases, for only five innovations did both cumulative normal and lognormal curves fail on this score (and for one of these - VD - the result became inconclusive when C_t was added to the equation.) Indeed, for only eight of the twenty-two was the model incapable of producing an equation for which autocorrelation could be conclusively ruled out (as opposed to nineteen for the logistic.)

A more stringent test, however, is whether the model can explain which innovations will have a cumulative normal and which, cumulative lognormal diffusion curves. It has been argued[1] that relatively cheap and simple innovations which can be produced off site (Group A) should be characterised by lognormal diffusion and more sophisticated, expensive innovations which must be at least partly built on the adopters' premises (Group B) by cumulative normal diffusion.

On the basis of the technical descriptions given in Appendix one, fairly unequivocally, 7 of the sample innovations fall into Group B and 8 into Group A. The remaining 7 innovations each share certain characteristics with both groups,[2] for instance, numerically controlled machine tools are technically complex, but relatively inexpensive and can be built off site, and they do not require lengthy installation periods. As it happens, these three groups could have been almost exactly predicted had the innovations been ranked according to cost, but then this is not surprising as technical complexity and physical size must largely determine cost.

Table 6.7.1. The cost of innovation groups A and B.

	No. of Innovations.	Range in cost of typical[i] initial outlay.
Group A	8	All £15,000 and below.
Group B	7	All £100,000 and above.
Unclassified	7	All £20,000 - £55,000.[ii]

i. typical figures taken from Appendix one, ideally to represent the cost of the average installation at the mid-point of the diffusion period.
ii. except for one innovation, S.L., which only cost £3,000 per loom. However looms are always installed in batches, therefore the typical installation cost of a batch of S.L's would probably lie in the £20,000 - £55,000 range.

for footnotes see next page -

1. See chapter 5 section 5 and chapter 3 section 7.
2. Of course, some ambiguities and overlaps are bound to occur in any simple two way classification; for terms such as technically complex and cost, there are no obvious cut-off levels below which an innovation may be defined as 'cheap' and 'simple.' There is, rather, a distribution of these characteristics and any innovation may be more or less as described by 'Group A' or 'Group B'.

Table 6.7.2. lists these groupings and pools the results of the three previous tables of results. Generally, expectations as to the most appropriate curve are fulfilled. For all but one (CC) of the Group B innovations, a cumulative normal diffusion curve is most appropriate; as predicted by the model. Moreover, except for CC the recorded $\bar{R}^2$ exceeds .9 in every case, and for all but CC and VD, the hypothesis of significant positive autocorrelation may be rejected.[1] Even for VD, when C_t is added to the equation, whilst it does not appear as a significant influence (see table 6.6.1), its presence is sufficient to raise the D-W statistic from the range indicating significant autocorrelation into the inconclusive range.

Similarly, for all but one (WSB) of the Group A innovations, a cumulative lognormal diffusion curve is most appropriate - again as predicted by the model. Only 2 of these 8, WSB and EH, exhibit significant autocorrelation.

Interestingly, whilst each curve performs well for its 'own group' of innovations, neither of them is as impressive in explaining diffusion for the 'other's group'. Thus the cumulative lognormal provides an acceptable explanation of the diffusion for only four Group B innovations and, for three of these, the $\bar{R}^2$ are considerably lower than those recorded for the cumulative normal.

This is a fairly encouraging result in that it indicates that the data is precise enough to allow a differentiation between the alternative hypotheses. There is always a danger with diffusion data that <u>any</u> upward sloping time trend will provide an adequate explanation, regardless of the shape of the true

1. Only those regressions reported in table 6.4.1 and table 6.5.1, and the regressions in 6.6.1 with significant C_t are considered in this assessment.

relationship. In this case, however, at least for the obvious Group A and B innovations, this problem has not materialised.[1]

It is not surprising that the picture is not so clear cut for the unclassified innovations. There is not sufficient technical information to be able to decide beforehand which of the two types each one of these innovations should most resemble. Generally, one might expect neither the normal nor the lognormal to perform so well for this group as they did for B and A respectively, rather, it might be expected that the two curves will be closer together in their explanatory performance with respect to individual innovations. As can be seen from the table, both of these expectations are largely borne out. The cumulative lognormal provides the better fit for two and the cumulative normal for four innovations (for CT the fits are almost identical.) Secondly, for three of these innovations (TC, NCPP and CT), <u>both</u> curves yield acceptable fits and for three of the other four, <u>neither</u> curve provides an acceptable fit; therefore, for only one unclassified innovation (NCTB) does one of the alternative curves provide a fit which is both acceptable and obviously superior to the fit offered by the other curve.

If anything, the cumulative normal performs rather better for this group as a whole: it provides a better fit than the lognormal in four of the seven innovations, an almost identical fit for one and an acceptable fit for the two cases in which the lognormal yields the better explanation. This conclusion must be seen in the light of the discussion of the previous paragraph, however.

Finally, what are the reasons for the model's failure to adequately explain the diffusion of certain innovations? From table 6.7.2, there are five innovations for which an acceptable fit was not obtained in any of the regressions: VD, EH, SL, SPC and NCTURN. In each case $\bar{R}^2$ was fairly high (particularly so for VD, EH and SL) but the recorded D-W indicated conclusively the presence of significant positive autocorrelation. Now the presence of

1. The generally poor performance of the logistic also tends to re-inforce this particular conclusion.

Table 6.7.2. The most appropriate time curve for each innovation.

Group B Innovations.	(a) most appropriate curve.	$\bar{R}^2$	D-W	(b) Comments.	$\bar{R}^2$	D-W
BOP	C.Norm	.875	1.946	C.Lognorm acceptable.	.815	1.358
CC	C.Lognorm.	.792	1.068(*)			
ATL	C.Norm.	.924	1.786	C.Lognorm acceptable.	.845	(*) 1.251
TK	C.Norm.	.978	1.647			
VD	C.Norm.(but unacceptable fit)	.937	.953 *	If insignificant C_t included, acceptable.	.940	(*) 1.225
VM	C.Norm.	.906	1.410			
PCBC	C.Norm.	.951	1.551	C.Lognorm acceptable.	.784	1.830
Group A Innovations.						
SP	C.Lognorm.	.926	1.663			
WSB	C.Norm.	.904	1.296(*)			
ASB	C.Lognorm.	.983	2.089			
F	C.Lognorm.	.968	2.368	C.Norm acceptable	.943	1.302
SF	C.Lognorm.	.986	1.904			
GA	C.Lognorm.	.959	2.165	C.Norm acceptable	.940	1.673
ADH	C.Lognorm.	.816	2.150			
EH	C.Lognorm.(but unacceptable fit)	.971	.877 *			
Unclassified.						
TC	C.Lognorm.	.816	1.467	C.Norm acceptable	.708	(*) 1.261
NCPP	C.Lognorm.	.979	3.358	C.Norm acceptable	.865	(*) 1.201
CT	C.Lognorm.	.892	1.695			
	C.Norm.	.892	1.982			
NCTB	C.Norm.	.906	1.147 (*)			
SL	C.Norm.(but unacceptable fit)	.944	.334 *			
SPC	C.Norm.(but unacceptable fit)	.868	.681 *			
NCTURN	C.Norm.(but unacceptable fit)	.865	.475 *			

Notes: 1. The most appropriate curve is defined, for each innovation, as the one which achieves the highest $\bar{R}^2$ and D-W. See tables 6.4.1, 6.5.1. and 6.6.1; however only those regressions with $\hat{c}$ significantly different from 0 are considered.
2. An unacceptable fit is defined as one for which significant autocorrelation is present and/or $\bar{R}^2 < .75$.
3. (*) and * as defined in table 6.4.1.

autocorrelation must always drastically reduce the usefulness of the equation concerned,[1] but in this case the problem is particularly damning if it can be shown to be the result of mis-specification of mathematical form: after all, the main point of the exercise is to test the predictions of the model about the true mathematical forms of the diffusion curve.

In the case of VD, as has been mentioned, when C_t is added to the cumulative normal curve, the D-W statistic moves into the inconclusive range, indicating that there is no longer conclusive proof of serial correlation in the theoretical disturbance term. This suggests that the apparent autocorrelation for the simple cumulative normal may have been due to an omission of the cyclical influence on $\bar{Q}_t$, rather than a fundamental mis-specification of mathematical form. The fact that C_t, as specified, is not significant, when included, may simply indicate that it does not accurately represent the cyclical influences on $\bar{Q}_t$.

For EH, there is no evidence to suggest that omission of a cyclical influence is the reason for autocorrelation, a visual inspection of the residual plot for the cumulative lognormal equation indicates under-prediction in the early years of diffusion, followed by consistent over-predictions for the rest of the curve. Thus, whilst the residuals are small in magnitude ($\bar{R}^2 = .971$), they are quite obviously serially correlated - quite possibly because of mis-specification of form. As can be seen from table 6.4.1, the cumulative normal seems to be even less appropriate than the lognormal: an exceptionally low D-W and relatively low $\bar{R}^2$ being recorded.

For SL, SPC and to a lesser extent, NCTURN, the serial correlation in the residuals might well be due to errors in the measurement of the dependent variables. As is explained in Appendix 2, diffusion data for these three innovations was computed from information from samples of firms in the industries concerned. Whilst the sample for NCTURN seemed to be fairly

1. Typically, inefficient predictions, underestimates of the standard errors and invalidation of t tests result.

representative,[1] for SL and SPC they were certainly not random but rather rough and ready. Assuming that biassed estimates of the population $\bar{Q}_t$ resulted from these poor samples, and that the bias was pervasive over time, this could quite possibly produce serial correlation in the disturbance term. Of course, one cannot rule out mis-specification of form as an added possibility, but it is notable that SPC and SL are the only two innovations for which serious doubts are justified as to the adequacy of the samples from which the diffusion data has been computed.

Given that autocorrelation is apparent using either curve for these five innovations, the obvious question is does this necessarily reduce the information value of the estimates $\hat{b}$ and $\hat{a}$ in these cases? After all autocorrelation, of itself, does not produce biassed estimates. However, the factors leading to autocorrelation may also produce bias. This is certainly true for mis-specification of mathematical form which leads not only to serial correlation of the disturbance term but also to correlation between the disturbance term and the explanatory variable, t, and in turn, therefore, to biassed $\hat{b}$. On the other hand, neither measurement errors in $\bar{Q}_t$ nor the omission of the cyclical variable, C_t, should lead to bias as neither the measurement errors nor C_t should be correlated with time.

In other words, for EH and perhaps NCTURN, the estimates of $\hat{b}$ and $\hat{a}$ are quite possibly worthless, but for VD, SPC and SL there is sufficient reason for retaining them, at least initially, in the set of diffusion parameters to be used in the cross-section analysis experiments in chapter 8.

1. As was true for the other handful of innovations for which data was computed from samples of firms. See Appendix 2.

Summary.

The results of this chapter have been fairly encouraging. Both the simple cumulative normal and lognormal curves offer a better explanation of the diffusion of the 22 sample innovations than does the logistic. Taken together, as the model suggests they should be, their superiority is even more pronounced. All but five of the innovations' diffusion curves can be adequately explained by either the normal or the lognormal and for only two of these five does the evidence suggest that an alternative trend curve would necessarily have offered a better explanation. Perhaps more importantly, the model has been fairly successful in predicting which innovations will have lognormal and which normal diffusion curves.

On the other hand, the introduction of the cyclical variable, C_t, into the analysis has been only partially justified by the results reported in table (6.6.1). In some cases inclusion of this variable does produce an improved fit, but for the majority of innovations there is no evidence that diffusion is sensitive to the cyclical influence suggested by the model. However, the rather poor quality of data available for measurement of C_t makes this particular conclusion very tentative.

Appendix One to Chapter 6 : Changing industry size.

This appendix reports an attempt to broaden the empirics, by allowing for the effects of changes in the sizes of the industries in the sample.

The results reported in this chapter are based on the assumptions that μ_{st} and σ^2_{st} are invariant with respect to time. While there is no evidence to suggest that σ^2_{st} has changed radically in any of the industries over the diffusion periods[1], it would be surprising if the overall size of these industries has remained constant. Given a consistent lognormal distribution of firm size and constant σ^2_s, a change in the overall size of an industry must be reflected in equivalent changes in μ_{st}.

If S_j, n_j and $\bar{S}_j$ denote, respectively, the aggregate industry size, the number of firms and arithmetic mean firm size in industry j, then $S_j = n_j \bar{S}_j$.

As the arithmetic mean of a lognormal distribution is given[2] by

$$e^{\mu_s + \frac{1}{2}\sigma^2_s}, \text{ then}$$

$$\log S_j = \log n_j + \mu_{sj} + (1/2)\sigma^2_{sj}$$

and if n_j and σ^2_{sj} are both constant over time[3], any changes in industry size must be reflected by changes in μ_{sj}.

1. Indeed, past research suggests that σ_s^2 may be, generally, a fairly stable parameter. See footnote 1, page 19 of the previous chapter.
2. See Aitchison and Brown, op.cit., p.12.
3. In the two sample industries in which new entry and deaths have been significant in the periods considered, diffusion has been studied only for those firms existing at both the start and end of the time period. Therefore n_j should also be stable.

Unfortunately, annual time series data is generally unavailable for S_j for most industries, mainly because the level of aggregation required in this study is much lower than that for which official statistics are provided. As a second best solution, the level of S at the start and end of the periods considered have been used to calculate trend growth rates of the form $S_{jt} = S_{jo}e^{\delta_j t}$.

S has been measured by industry employment, the source being the five yearly Census of Production reports for the years nearest to the dates required. Even then, however, the MLH classifications used in the Census by no means always conform to the industries considered here. (For instance, the Printing Press manufacturing industry in which numerical control is diffusing - NCPP - is only one <u>very</u> small part of the MLH 'Other mechanical engineering.')

Nevertheless, estimates of δ_j were made for all j. These estimates were then used in a new set of curve fitting exercises using the analysis in appendix 2 to the last chapter.

From equations (5.A2.6) and (5.A2.7) if $S_{jt} = S_{jo}\, e^{\delta_j t}$ and

$\mu_{stj} = \mu_{soj} + \delta_j t$,

then $\quad \bar{Q}_{jt} = N\,(\Psi_j \log t + \beta_j \delta_j t + \Omega_j\, C_{jt} \mid - \log\alpha_j - \beta\mu_{soj}, \beta_j^2 6_{sj}^2 + 6_j^2)$

or $\quad \bar{Q}_{jt} = N\,(\Psi_j t + \beta_j \delta_j t + \Omega_j C_{jt} \mid - \log\alpha_j - \beta\mu_{soj}, \beta_j^2 6_{sj}^2 + 6_j^2)$

The inclusion of the extra term, $\beta\delta t$, leads to new estimating equations (in place of 6.1.4a and 6.1.4b):

$$\bar{Z}_{jt} = a_j + b_j \log t + c_j C_{tj} + d_j t \qquad (6.A.1)$$

or

$$\bar{Z}_{jt} = a_j + b_j t + c_j C_{tj} + d_j t \qquad (6.A.2)$$

where

$$d_j = \frac{\beta_j \delta_j}{(6_j^2 + \beta_j^2 6_{sj}^2)^{\frac{1}{2}}} = \frac{\delta_j}{\left(\frac{6_j^2}{\beta_j^2} + 6_{sj}^2\right)^{\frac{1}{2}}} \qquad (6.A.3)$$

Estimation of equation (6.A.1) would be almost certainly pointless, however, as it is probable that $\hat{b}$ and $\hat{d}$ would be imprecise due to collinearity between t and log t. Therefore the best course of action to take seemed to be to impose $\hat{d}$ on the equations as follows:

$$\bar{Z}_{jt} - \hat{d}_j t = a_j + b_j \log t + c_j C_{jt} \tag{6.A.4}$$

$$\bar{Z}_{jt} - \hat{d}_j t = a_j + b_j t + c_j C_{jt} \tag{6.A.5}$$

This required a prior knowledge of $\hat{d}_j$ of course, but a method is available for deriving estimates independently of the time series data. The three constituent parts of d_j in (6.A.3) are δ_j, 6^2_{sj} and $(6/\beta)^2_j$. Estimates of δ_j were derived as above, $\hat{6}^2_{sj}$ were available from Appendix 5 and $(\widehat{6/\beta_j})$ were computed by fitting Quasi-Engel curves for each innovation. (As this forms an integral part of the empirics reported in the next chapter, the discussion of the exact method used is postponed until then.)

Thus it was possible to compute the new dependent variable $(\bar{Z}_{tj} - \hat{d}_j t)$ for all of the sample innovations. The regressions reported in the main text of this chapter were then re-computed using this new variable.

Results were, generally, rather disappointing. For the cumulative normal, of course, $\bar{R}^2$, D-W, $\hat{a}_j$ and $\hat{c}_j$ all remained unchanged and, in every case, $\hat{b}_j$ was reduced by an amount exactly equal to $\hat{d}_j$. Indeed, it was unnecessary to recompute these regressions as the new $\hat{b}_j$ must be simply the difference between the old $\hat{b}_j$ and $\hat{d}_j$. For the cumulative lognormal, there were very small changes in all test statistics and coefficients but only for $\hat{b}_j$ did these amount to more than a change in the second decimal place. Again, $\hat{b}_j$ were reduced, but in this case, of course, the new $\hat{b}_j$ could not be predicted from a knowledge of the old $\hat{b}_j$ and $\hat{d}_j$.

Overall, none of the conclusions of table 6.7.2 are altered. Moreover, because of the, generally, very low values of δ_j and thus d_j (a typical value for $(6^2_{sj} + (6/\beta)^2_j)^{\frac{1}{2}}$ is about 3 to 4), changes in $\hat{b}_j$ for both cumulative normal and lognormal curves were also very small. More importantly, their estimated variance across innovations was virtually unchanged.

Consequently, the cross-section analysis of chapter 8, which attempts to explain inter-industry and innovation differences in $\hat{a}_j$ and $\hat{b}_j$ are based on the estimates reported in the main text. It seems probable that much the same results would ensue using the estimates derived from the above alternative regressions. At any event, the data problems mentioned are sufficient reason for doubting the value of these new estimates.

Chapter 7 : Empirical Quasi-Engel curves.

This chapter provides an alternative test of the model using cross-section data for each innovation. In addition, however, it yields a further set of observations across industries for two of the parameters of the model. In the next chapter, a cross-industry explanation will be attempted for both the diffusion curve and Quasi-Engel curve parameters.

In section 3 of chapter 5, it was shown that, at time t, the probability of having adopted is related to firm size by an upward sloping S shaped curve - the cumulative lognormal. The mathematical formulation of this relationship is:

$$P\left\{q_{it} = 1 \,\middle|\, S_{it}\right\} = \Lambda\left(S_i \,\middle|\, \frac{-\log\alpha - \log\theta_t}{\beta}, \sigma^2/\beta^2 \right) \qquad (5.3.8.)$$

Given the assumptions made about σ^2, β^2 and θ_t, this implies that the variance of this distribution is constant over time, but that the mean decreases over time. Thus the probability of adoption rises monotonically with time for all firms. As the variance does not depend on the specification of θ_t, Group A and Group B innovations will differ, in this context, only in the way in which their Quasi-Engel curves shift to the left over time.

Section 1 of this chapter explains the empirical method used to test these predictions and discusses the problems resulting from the limited quantity of data available. Section 2 presents the results obtained and discusses their implications.

1. The empirical method and data limitations.

Denoting the conditional probability of a firm of size S_i having adopted by time t as P_{it}, (5.3.7.) may be expressed, quite generally, as:

$$P_{it} = P\left\{q_{it} \,\middle|\, S_{it}\right\} = N\left(\log S_{it} \,\middle|\, \frac{-\log\alpha - \log\theta_t}{\beta}, \sigma^2/\beta^2 \right) \qquad (7.1.1.)$$

$$= N\left(\frac{\beta\log S_{it} + \log\alpha + \log\theta_t}{\sigma} \,\middle|\, 0, 1 \right) \qquad (7.1.2.)$$

$$\text{If} \quad z^1_{it} = a^1 + b^1 \log S_{it} \quad \text{where } a^1 = \frac{\log\alpha + \log\theta_t}{6} \text{ and } b^1 = \beta/6 \quad (7.1.3)$$

then $P_{it} = N\,(z^1_{it}|0,1)$

That is the normit, or normal equivalent deviate, of the probability is a linear function of the logarithm of the firm's size. But to make (7.1.3) operational requires an empirical measure of the probability itself. As with the diffusion curve, the probability is replaced by an observed frequency. Following the convention of Quasi-Engel curves, the population is split into a number of size ranges and P_t for each range is measured by $\bar{P}_{kt}$ (the observed proportion of firms in the kth range having adopted) and S_{kt} is measured by the mean size of firms in the range. So long as the ranges are not too wide, this should provide an adequate approximation.[1]

Therefore, the empirical counterpart of (7.1.3) is given by:

$$\bar{z}^1_{kt} = a^1 + b^1 \log S_{kt} \qquad (7.1.4)$$

where $\bar{z}^1_{kt}$ is the normit of the observed penetration (frequency of adopters) in the kth size range at time t.

Because observed frequencies have been substituted for probabilities, as for the diffusion curves, the theoretical disturbance term will not be homoscedastic and the choice of estimating technique again rests between minimum normit and maximum likelihood.[2]

However, as can be seen from appendix 4, the number of firms for which data is available in most industries is relatively small (usually less than 50). In order for the penetration rates of size classes not to be too sensitive to the behaviour of individual firms, it is considered advisable that each size class should include at least eight firms. Further, for some industries, the only data available is already grouped into size classes, and usually only four classes at that. The unfortunate upshot of this is

1. By fitting the curve to class or range means, all firms in any range are treated as if they had the same size - the class mean.

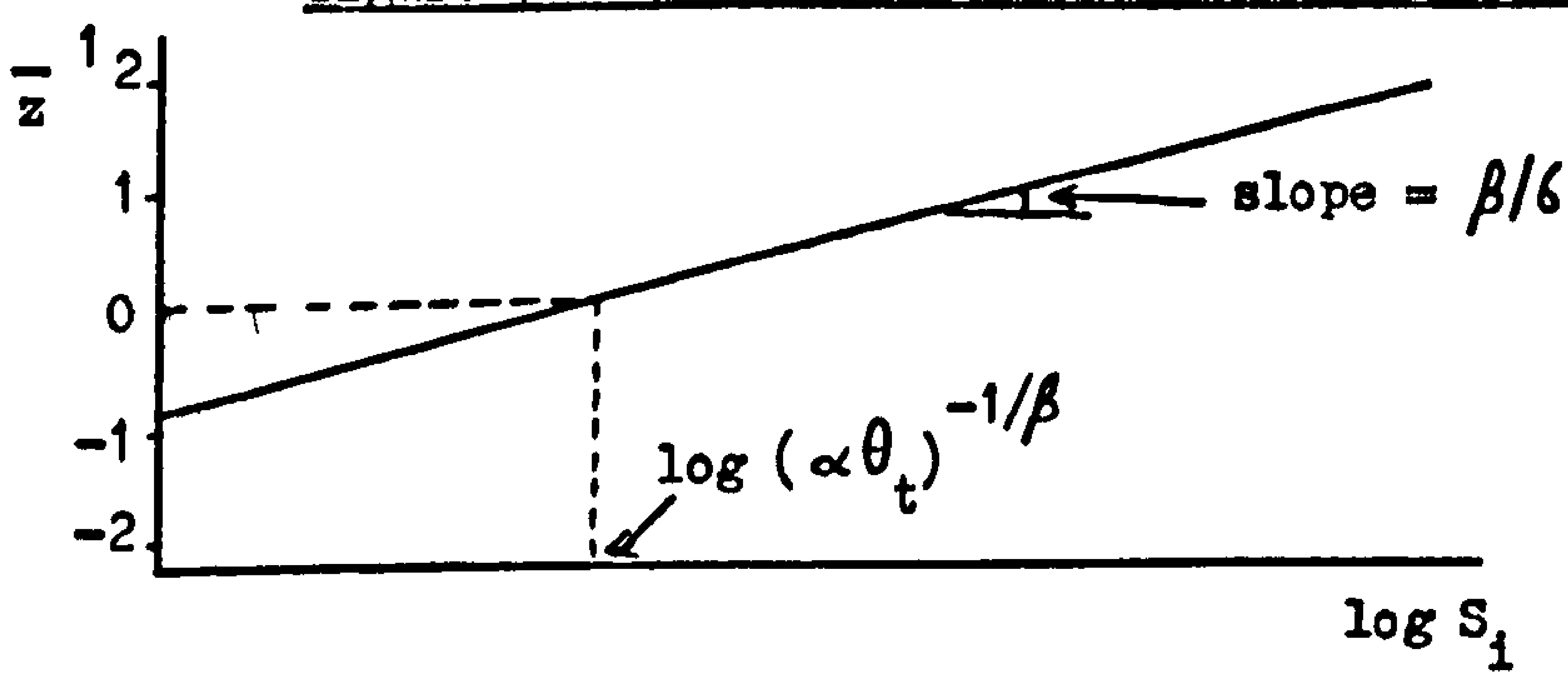
2. The argument is exactly equivalent to that already described in section 2 of chapter 6.

that on average, only four observations are available for each industry.

Quite obviously it would be pointless to use sophisticated empirical techniques under such circumstances. Instead, a purely graphical method[1] is used which amounts, quite simply, to drawing a straight line through the observations plotted on a graph, with the horizontal axis measuring the logarithm of size and the vertical axis the normits of $\overline{P}_{kt}$. The fitted line has a slope, $\beta/6$, equal to the inverse of the standard deviation of the critical size distribution and cuts the line: $\overline{z}^{\,1}_{kt} = 0$ at:

$$\log S = -\left(\frac{\log \alpha + \log \theta_t}{\beta}\right),$$ that is the mean of the critical size distribution.

figure 7.1.1. The transformed lognormal Quasi-Engel curve.



N.B. This is equivalent to logarithmic probability paper[2] with $\overline{P}$ and S measured along the axes.

As a rough way of incorporating the weighting scheme implied by the more sophisticated statistical techniques, it is suggested[3] that the observations closest to $\overline{P} = .5$ ($\overline{z}^{1} = 0$) should be more important to the drawn line than are those close to $\overline{P} = 1$ or 0 ($\overline{z}^{1} = \pm\infty$).

To quote Aithison and Brown: 'Although (this graphical method) can hardly be regarded as a rigorous statistical test of lognormality, it provides a quick method of judging whether the population may be feasibly lognormal. Moreover, the parameters (of the distribution) may be estimated.'

1. Suggested by Aitchison and Brown(op.cit.), pp.31-34.
2. ibid. pp.31-33.
3. ibid. and also by Finney (op.cit.)

For each innovation it should be possible to fit the Engel curve at any and every point in time, and whilst the estimate of $-\left(\frac{\log\alpha + \log\theta_t}{\beta}\right)$ will vary over time, the slope[1] of the straight line (β/σ) should be insensitive to the point in time considered. Again, however, data limitations severely restrict the options available. For some innovations, this sort of cross-section data is available for only one point in time, for others, whilst penetration data is available for a number of alternative years, information on firm size is not. Moreover, in a few cases in the later years, penetration in the largest size class reaches 100%; as the normit of 100% is plus infinity, it is impossible to draw the Engel curve under these circumstances.[2] Therefore, the Quasi-Engel curves have been calculated at only one point in time for each innovation, and where there is a choice of date, the most recent year is chosen.

As it happens, this is hardly a serious limitation. If the lognormal hypothesis is true at all, it must be true for every point in time and, similarly, given the assumptions of the model, the variance of the distribution will be the same at every point in time.[3]

2. Results and implications.

The data used to compute the empirical curves is presented in Appendix 4. For each innovation the normit of penetration is plotted against the logarithm of mean size in each size class as in figures 7.2.1.[4] From each diagram, the mean and variance of the critical size distribution, (5.3.7.), are

1. In terms of normits and log. size, not in terms of proportions and size.
2. Unless one is prepared to approximate 100% by 99.9% which has a normit of +3. This is, of course, one of the problems of the normit transformation: whilst a Quasi-Engel curve can predict very high penetration rates (in excess of 99%) it can never predict 100% (or 0%) penetration for any size class.
3. For three innovations, it is possible to estimate the curve for a number of alternative years. Typically, $(\widehat{\sigma/\beta})$ does not vary much between years. For example, for SP, $(\widehat{\sigma/\beta})$ is 1.7 in 1966, 1.825 in 1968 and 1.825 in 1970.
4. An alternative approach was pursued for ATL, see the note at the end of Appendix 4.

calculated and reported in table 7.2.1. These diagrams and estimates are useful in a number of respects, three of which relate directly to equation 5.2.7, which is, of course, quite crucial for the whole development of the model:

$$\left\{\frac{R^*}{(ER)_N}\right\}_{it} = \alpha S_{it}^{\beta}\,\theta_t \epsilon_{it} \qquad (5.2.7)$$

Only if the error term ϵ_i in this expression is lognormally distributed will the critical size distribution (eqn 5.3.7) also be lognormally distributed. Now, to the extent that the time series implications of the model as a whole have proved fairly accurate, there is already some indirect support for this assumption. A more specific test is possible however. For the critical size distribution to be lognormal, clearly, the transformed Quasi-Engel curves shown in the figures should be described by straight lines; any systematic divergence from straight lines would suggest an alternative specification for critical size and thus ϵ_i. Unfortunately, given the small numbers of observations on which these figures are based, no strong test of the linear hypothesis is possible; moreover, the deliberate intent of fitting these curves closely to the central points and attaching less weight to extreme values of $\bar{z}^{1}_{kt}$ will give the impression of a poorer fit than is actually achieved.[1]

Visually, the straight line provides a fairly acceptable fit for most innovations. For all but 4 of the 21 innovations, there is an obvious monotonic relationship between penetration and firm size. For 10 of the innovations (F, SF, PCBC, BOP, EH, ADH, ASB, NCTN, NCTB, SPC) the straight line provides a close fit[2] and for another 7 (SP, WSB, VD, VM, NCPP, SL, TC) a fairly acceptable approximation which may have been bettered by some other

1. For instance, an observed error of +.5 normits at $\bar{z}'_{kt}=0$ means an error in $\bar{P}_{kt}$ of .1915, but a similar error at $\bar{z}'_{kt}=-2$ translates into an error in $\bar{P}_{kt}$ of only .0166.

2. But for SPC and NCTB this is hardly surprising as there are only 2 observations for each!

sort of curve in some cases. However, for both CC and TK, one pair of observations show no increase as class size increases and for two others, CT and GA, one pair of observations show an actual decline as class size increases.

The shortage of observations rules out the possibility of strong overall conclusions one way or the other but it is fair to claim that these figures at the very least, do not exclude the lognormal hypothesis - often the largest residuals occur at relatively low penetration rates where random influences will cause large distortions in the observed $\bar{z}'_{kt}$.

For the record, the standard χ^2 test of normality has been used (table 7.2.1); as can be seen, for none of the innovations can the hypothesis of a lognormal Quasi-Engel curve be rejected at the 5% level.

Secondly, it is now possible to be definite about the sign of β in (5.2.7). As was shown in section 3 of chapter 5, the Quasi-Engel curve will be upward sloping for positive β and downward sloping for negative β. Fairly unequivocally, all of the empirical curves are upward sloping. Pooling all innovations, there are 51 pairs of adjacent observations of which only 2 (as reported above) show a decline in diffusion between classes as size class increases, and two others show no change as size class increases. Thus over 92% of the observations are consistent with the hypothesis of positive β for all S. Clearly, then, size does matter : large firms on average do have a higher probability of having adopted at all times. This result should not be interpreted as meaning that larger firms are more progressive, however. From equation (5.2.6), $\beta = \beta_3 - \beta_2 - \beta_1$, and from the evidence of chapter 3, section 6, there are good reasons for believing that $\beta_1 < 0$, that is, that new innovations are more profitable for large firms, for purely technical reasons. It would be necessary to show that $\beta + \beta_1$, is positive (namely that large firms are more favourably disposed to new innovations and are better at gathering and interpreting information), before claiming that large firms are more progressive. Unfortunately, there is no way of testing this hypothesis.

Third, again with respect to equation (5.2.7), it is now possible to ascertain the relative explanatory power of S_{it} vis-a-vis ϵ_{it}. Of course one might test this equation, for each industry, by regression, using the form:

$$\log \left(\frac{R^*}{ER_N}\right)_{it} = \log (\alpha\theta_t) + \beta \log S_{it} + \log \epsilon_{it} \qquad (7.2.1)$$

on a cross-section basis at time t with observations of $\left(\frac{R^*}{ER_N}\right)_{it}$ and S_{it} for all firms. However, R^* and $(ER)_N$ would be impossible to measure for most firms, even if they were willing to co-operate.

Supposing, for the moment, that this data were available for all firms and equation (7.2.1) was estimated using ordinary least squares, then the R^2 would be calculated , as usual, by:

$$R^2 = \frac{\Sigma \hat{y}_i^2}{\Sigma y_i^2}; \text{ where } y_i = \log\left(\frac{R^*}{ER_N}\right)_{it} - \overline{\log\left(\frac{R^*}{ER_N}\right)_t} = \hat{\beta}(\log S_{it} - \overline{\log S_t}) + \widehat{\log \epsilon_{it}} \qquad (7.2.2)$$

and $\hat{y}_i = \hat{\beta}(\log S_{it} - \overline{\log S_t})$

Using the symbols of chapter 5,

$$\sigma_{st}^2 = \frac{\Sigma(\log S_{it} - \overline{\log S_t})^2}{n}, \quad \hat{\sigma}_t^2 = \frac{\Sigma(\widehat{\log \epsilon_{it}})^2}{n} \qquad (7.2.3)$$

where n is the number of observations.

$$\text{Thus, } R^2 = \frac{\hat{\beta}^2 \sigma_{st}^2}{\hat{\beta}^2 \sigma_{st}^2 + \hat{\sigma}_t^2}$$

$$= \frac{1}{1 + (\widehat{\sigma/\beta})_t^2 \, 1/\sigma_{st}^2} \qquad (7.2.4)$$

As estimates of σ_{st}^2 are available for all industries[1] (assuming no significant change over time) and $(\widehat{\sigma/\beta})$ have been estimated from the figures, R^2 may be calculated for all industries for the years in which the empirical Quasi-Engel curves have been estimated.[2]

1. See Appendix 5.
2. Indeed, if β, σ^2 and σ_s^2 are all relatively constant over time, these R^2 should apply to all years.

Interestingly, therefore, it is possible to compute R^2 for regressions which have never been calculated, because there are no observations for the dependent variable. This result is not only of interest value, however; given the quite central place afforded to the firm size variable in this model, a test of its explanatory power is very welcome.

Column 5 of table 7.2.1 lists the R^2 (in percentage terms) that have been computed in this way. The results are surprisingly powerful: for 7 of the 22 innovations,over half of the variance in ($\frac{R^*}{ER_N}$) is explained by the inclusion of S as the only independent variable and for another 4, $\hat{R}^2 > .4$. Using an F test, all 22 calculated R^2 are significant, of course, and in only 4 cases does R^2 drop below .2. The biggest limitation on these 'back door' R^2s is of course, that they have been computed using estimates of $(\widehat{\sigma/\beta})$ which must be rough and ready; on the other hand, there is no reason to suspect that these estimates are systematically biassed downwards which would result in an upward bias of R^2.

On the face of it, then, these results do justify the central role attributed to firm size but having said that, there is substantial unexplained variance in ($\frac{R^*}{ER_N}$) in all industries, which in this model is reflected in the error term ϵ_{it}.[1]

The fourth main point in estimating the empirical Quasi-Engel curves rests with the estimates obtained for (σ/β). These may be used in a cross-section study to test certain hypotheses presented in section 6 of chapter 5 as to the causes of inter-industry and inter-innovation differences in σ_j and β_j. No attempt will be made to use the estimates of the means of the empirical Quasi-Engel curves; their interpretation depends on the forms of θ_t used and the point in time to which they refer, moreover they add no new information over and above that already available from the more reliable time series regressions of the previous chapter. For the record, estimates of $\frac{-\log\alpha - \log\theta}{\beta}$ are noted in column 3 of the table.

1. Clearly, any further development of this model must be concerned with attempts to identify the other variables whose effect is being 'picked up' at present by ϵ_{it}.

Table 7.2.1. Estimated Quasi-Engel Curve Parameters.

Innovation(j)	$(\frac{\hat{\delta}}{\beta})_j$	$\left(\frac{-\log\alpha - \log\theta_t}{\beta}\right)_j$	$t_j =$	χ^2	Explanatory power of $\alpha_j S_{ijt}^{\beta_j} \theta_{jt} \epsilon_{ij}$ (%)
SP	1.825	6.1	9	.236	46.9
F	1.800	6.25	5	.007	47.6
SF	1.425	6.7	5	.014	59.1
WSB	2.9	7.925	14	.335	25.9
PCBC	2.625	10.425	6	0	29.9
BOP	0.96	8.86	11	.001	55.4
CC	3.75	10.40	11	.217	9.0
VD	3.125	8.05	14	.807	13.3
VM	1.65	6.00	11	.092	39.0
EH	0.86	6.01	22	.798	68.7
ADH	1.15	6.35	9	2.360	55.2
ASB	2.45	9.45	6	.078	21.3
NCPP	2.25	6.00	9	.079	35.4
NCTURN	1.20	5.85	15	.009	48.8
NCTB	2.234	5.616	14	-	28.2
GA	3.35	2.225	9	.310	12.0
SL	1.687	6.00	16	.210	50.4
TK	1.212	4.862	19	2.351	65.2
SPC	0.645	7.945	11	-	62.8
CT	1.34	6.76	9	2.09	25.3
TC	2.375	4.80	12	.239	40.5
ATL*	2.81	-	-	-	11.2

Notes: $(\frac{\hat{\delta}}{\beta})_j$ and $\left(\frac{-\log\alpha - \log\theta_t}{\beta}\right)_j$ have been computed from figures 7.2.1 - 7.2.21

t_j denotes the number of years between innovation j's first introduction and the year for which j's Quasi-Engel curve has been estimated.

The sixth column denotes the R^2 from the hypothetical regressions (7.2.1)

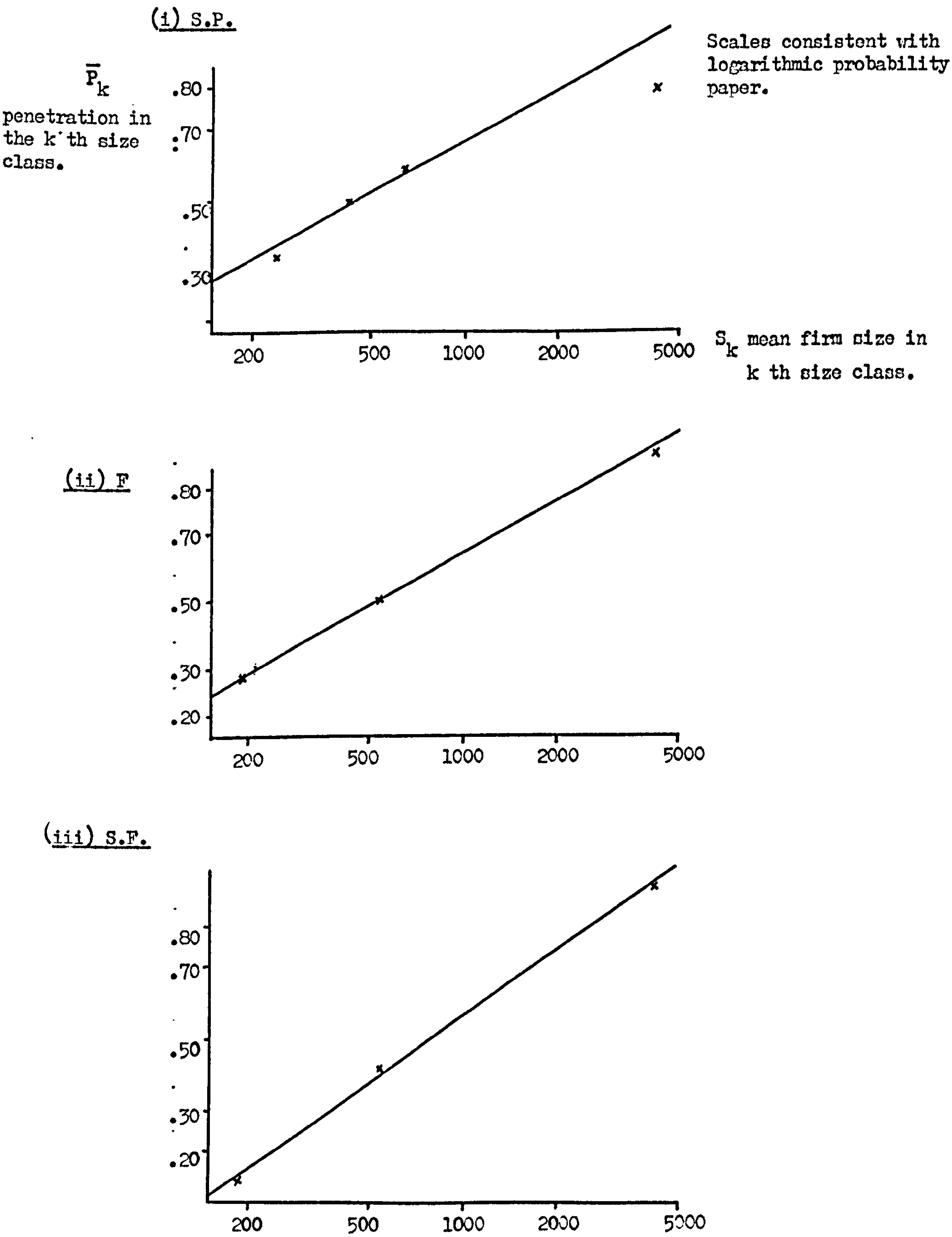
* For the method of calculating $(\widehat{\delta/\beta})$ for ATL, see Appendix 4.

The χ^2 have been calculated using the observed and predicted $\bar{P}'_{kt}$ transformed from the $\bar{z}'_{kt}$ in figure 7.2.1. None of these values is significant at the 5% level.

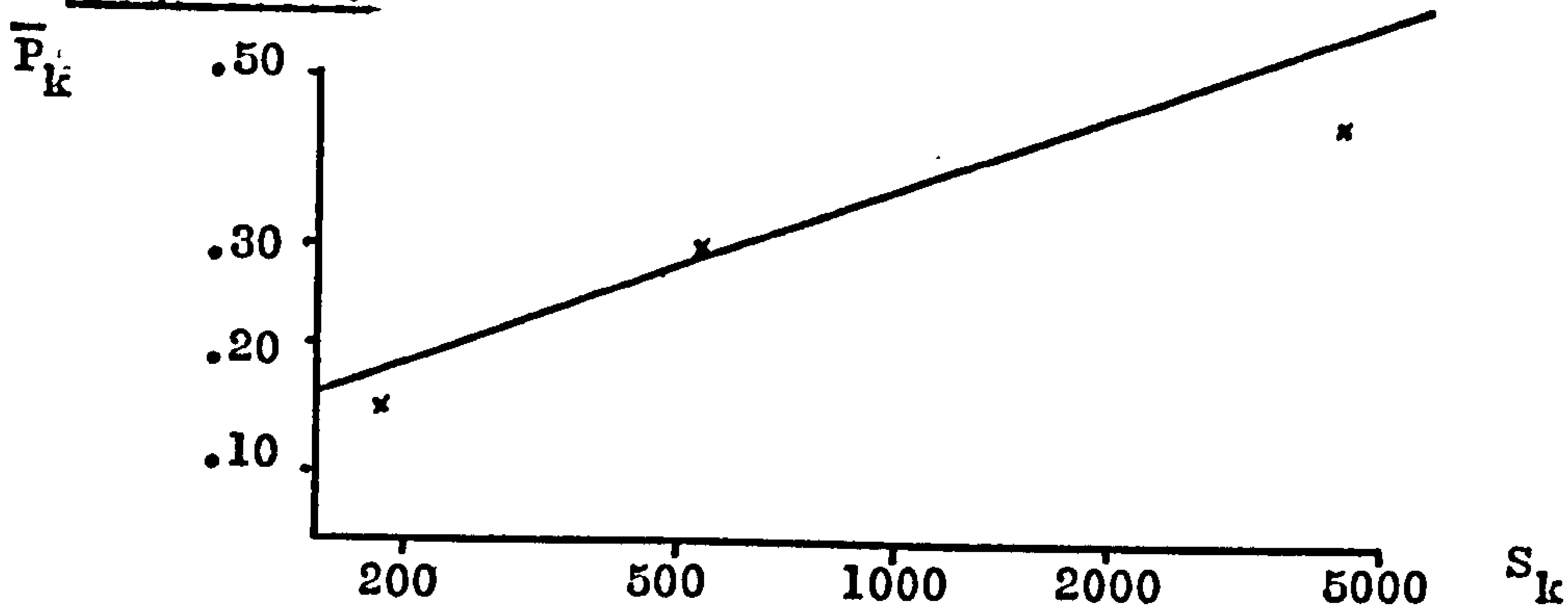
Summary

Perhaps the most important finding of this chapter is that, for all innovations, firm size has an important role to play in the diffusion process as indicated by column 5 of table 7.2.1. The estimated Quasi-Engel curves suggest quite strongly that β is always positive and that the assumption of a lognormal distribution for critical size may be reasonable. (The lack of extensive data requires that the latter conclusion be only tentative.) For each innovation an estimate of (σ/β) has been calculated which will be used alongside the estimated time series parameters from the previous chapter in the cross-industry analysis of chapter 8.

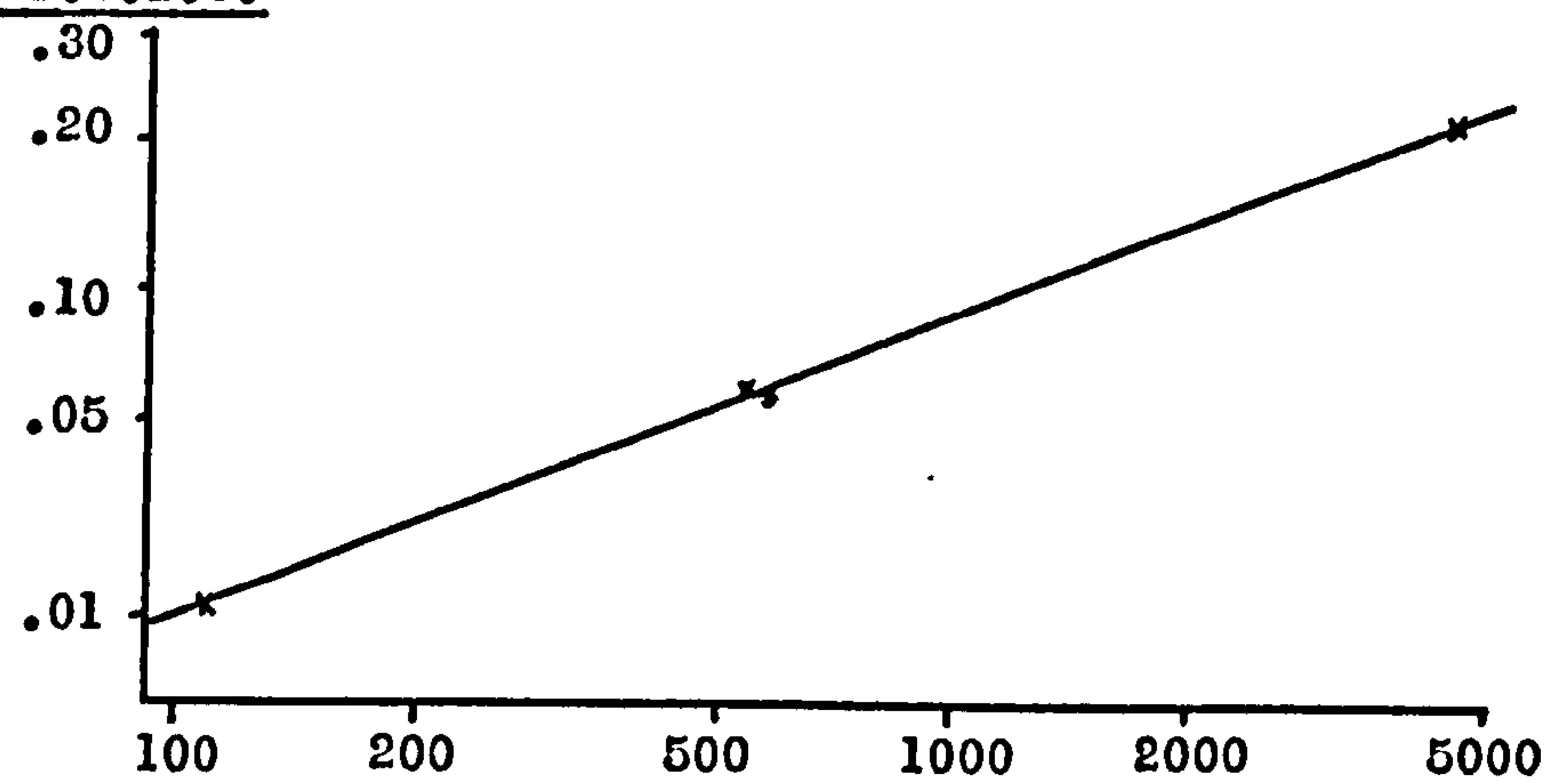
Figure 7.2.1. The empirical Quasi-Engel curves.



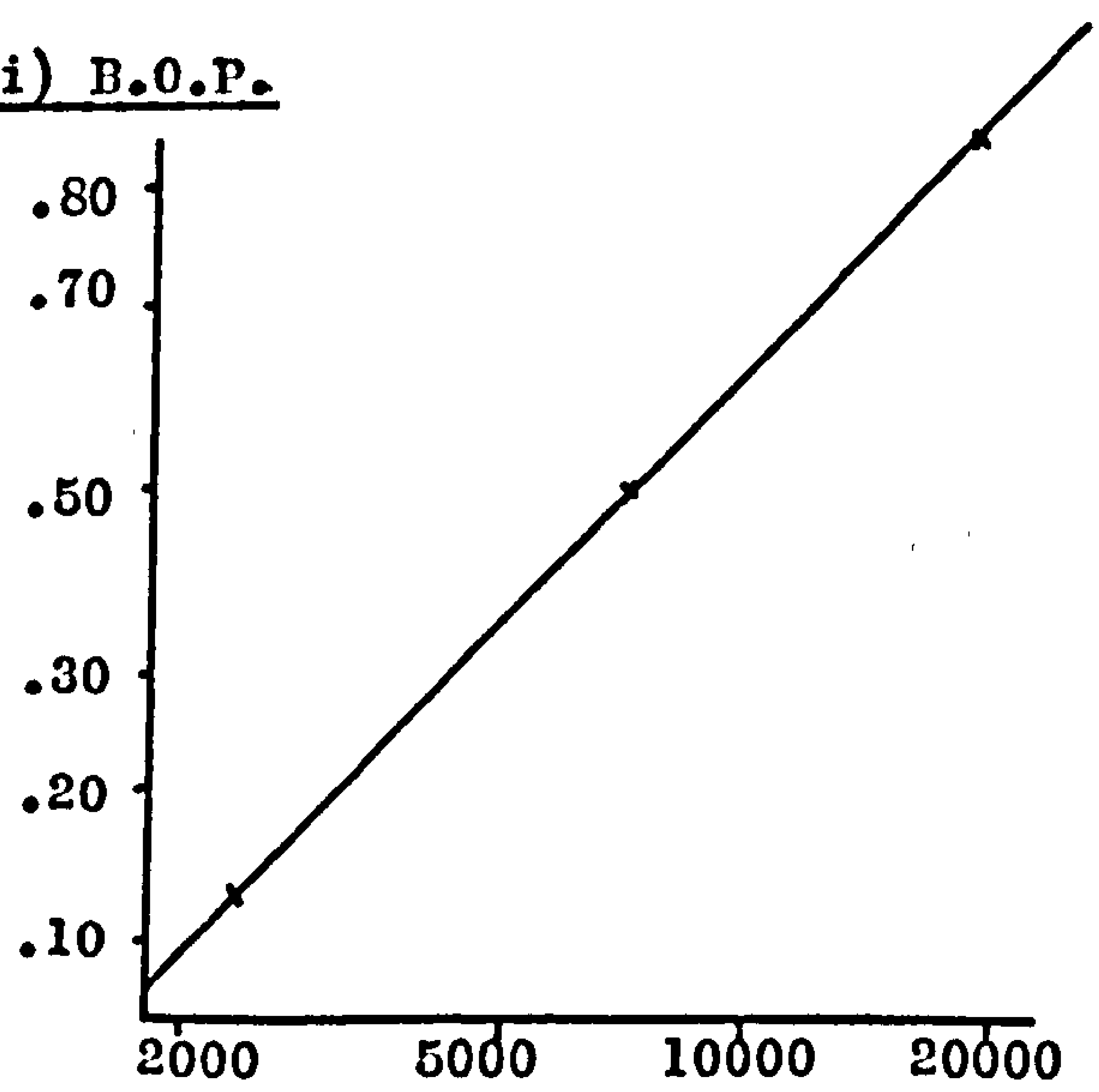
(iv) W.S.B.



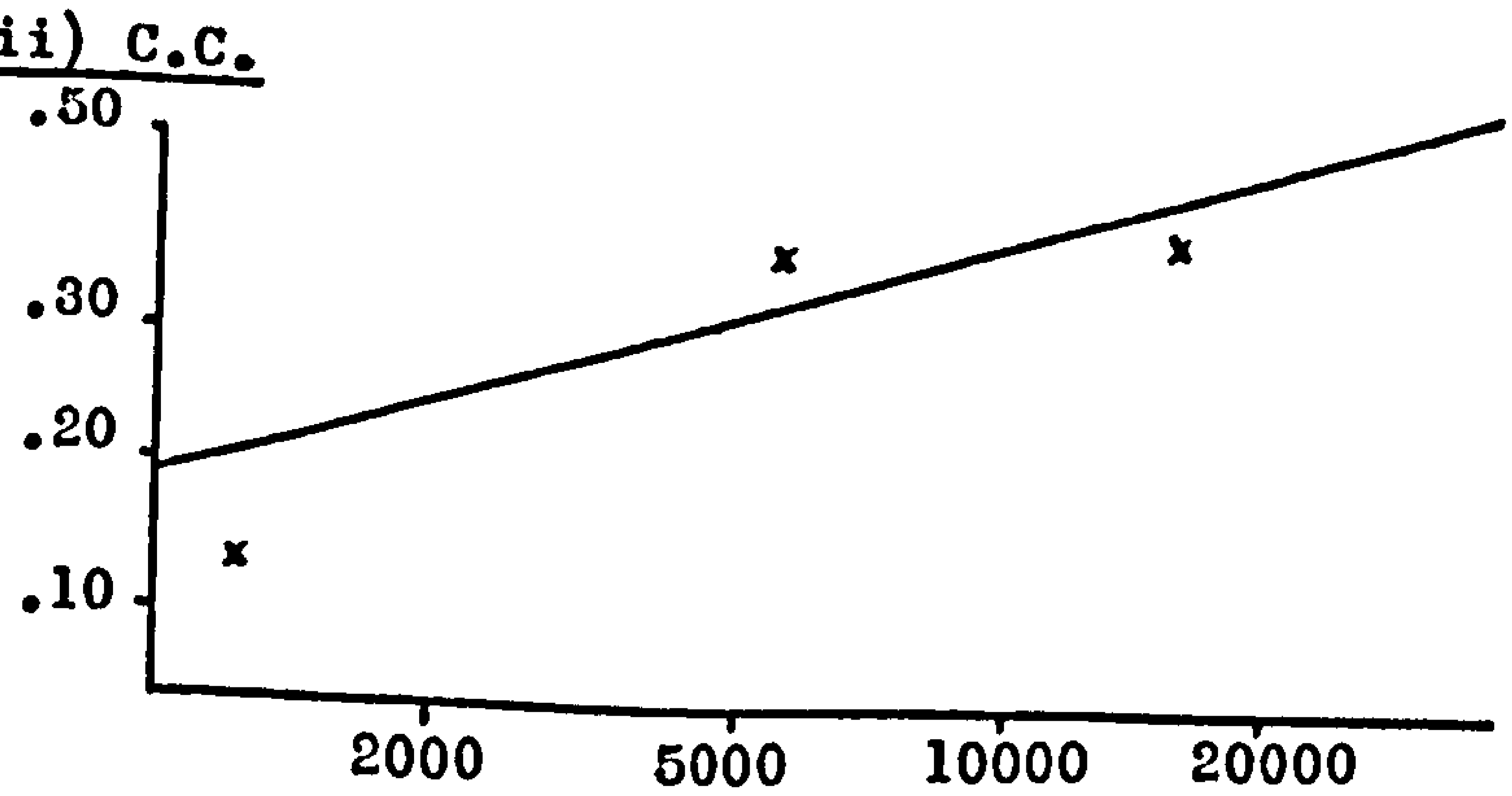
(v) P.C.B.C.

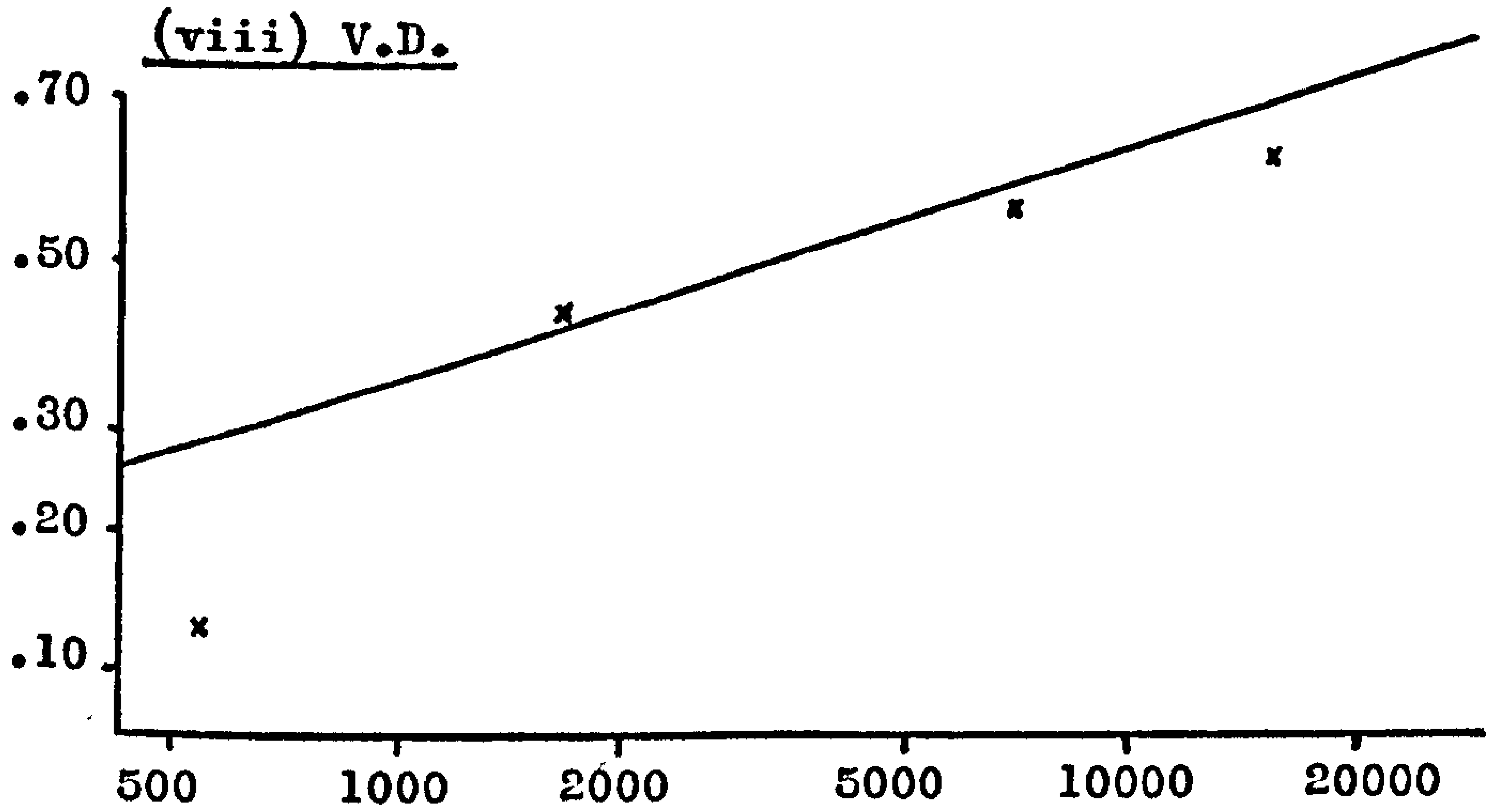


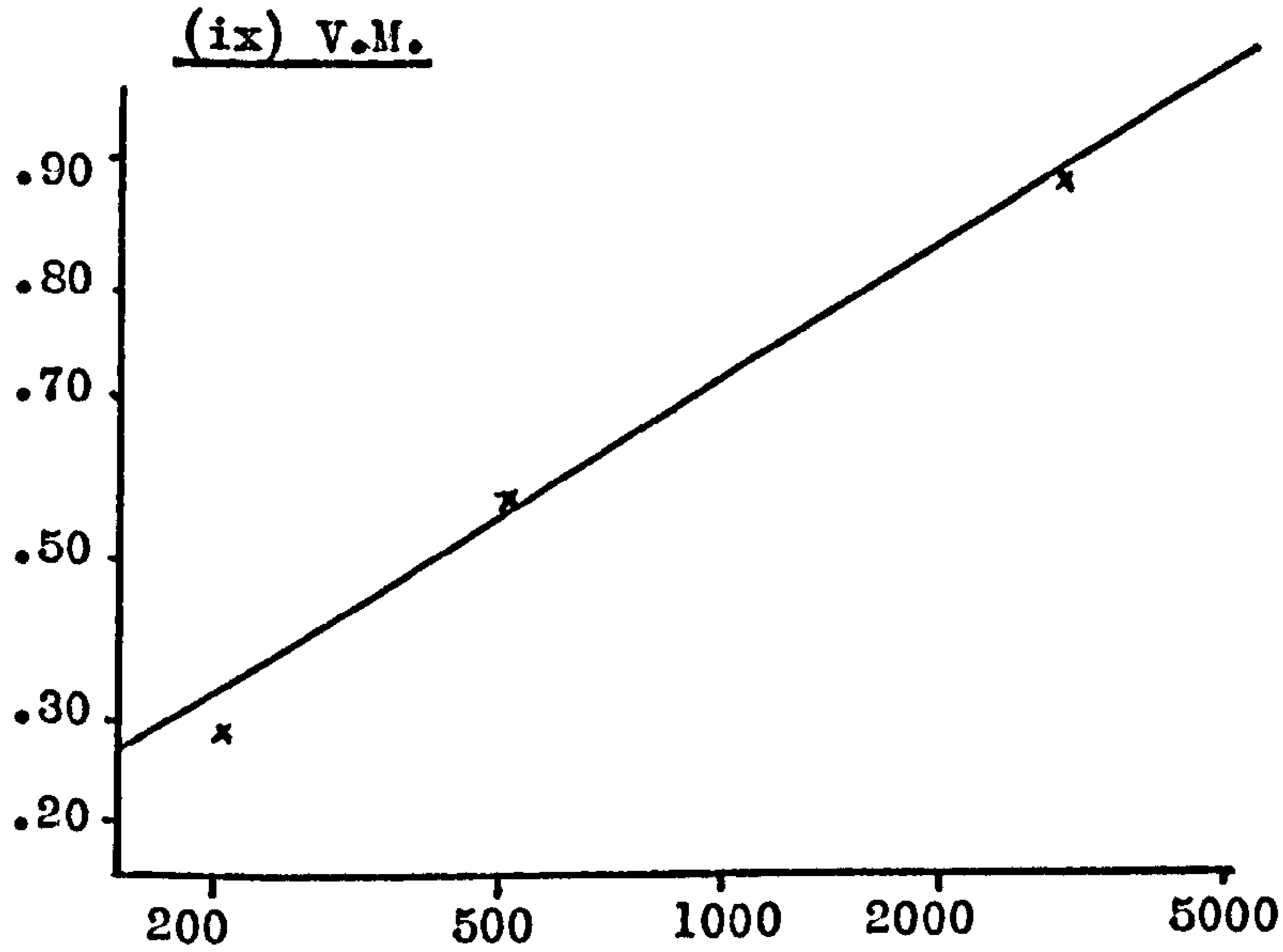
(vi) B.O.P.

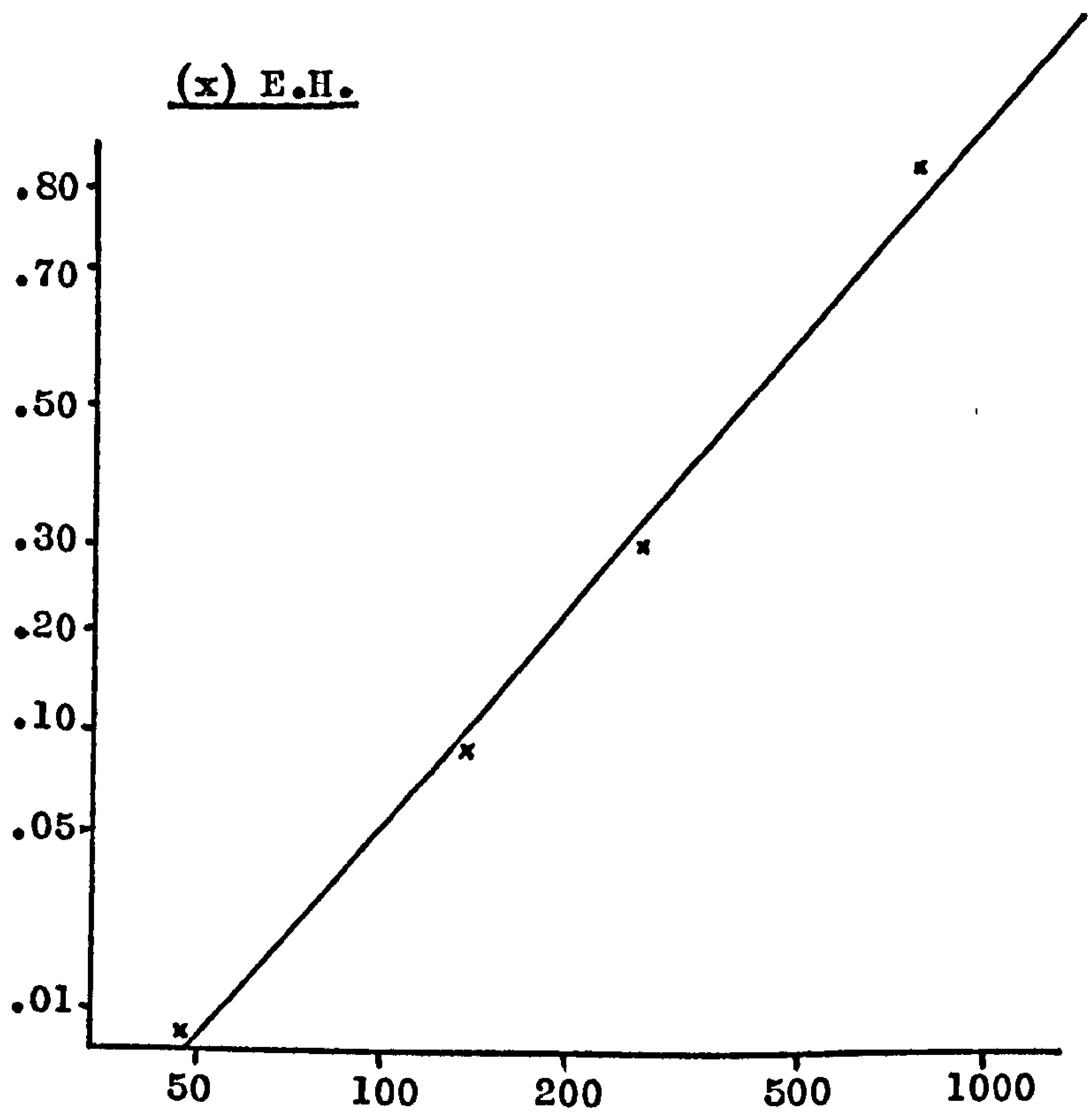


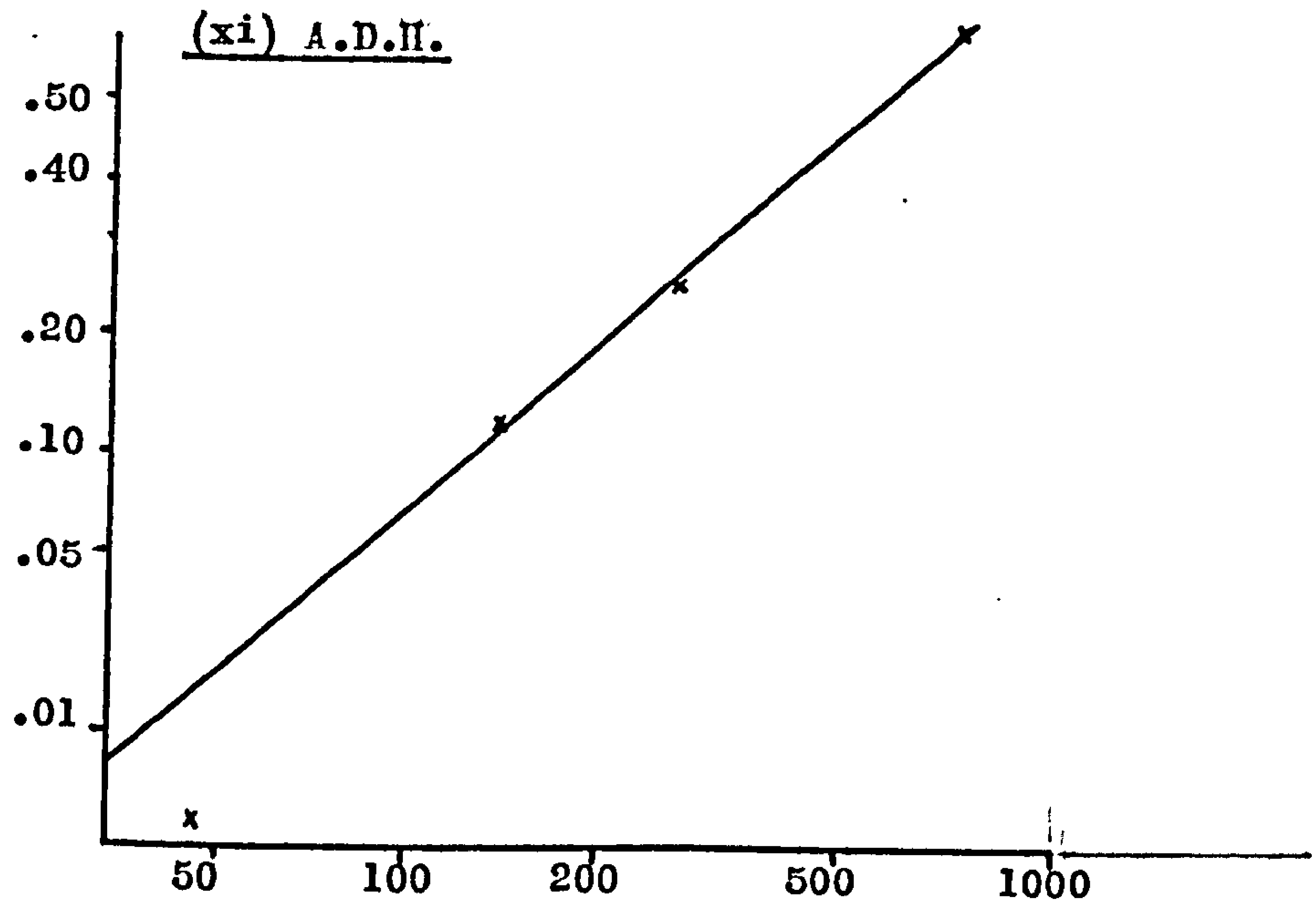
(vii) C.C.

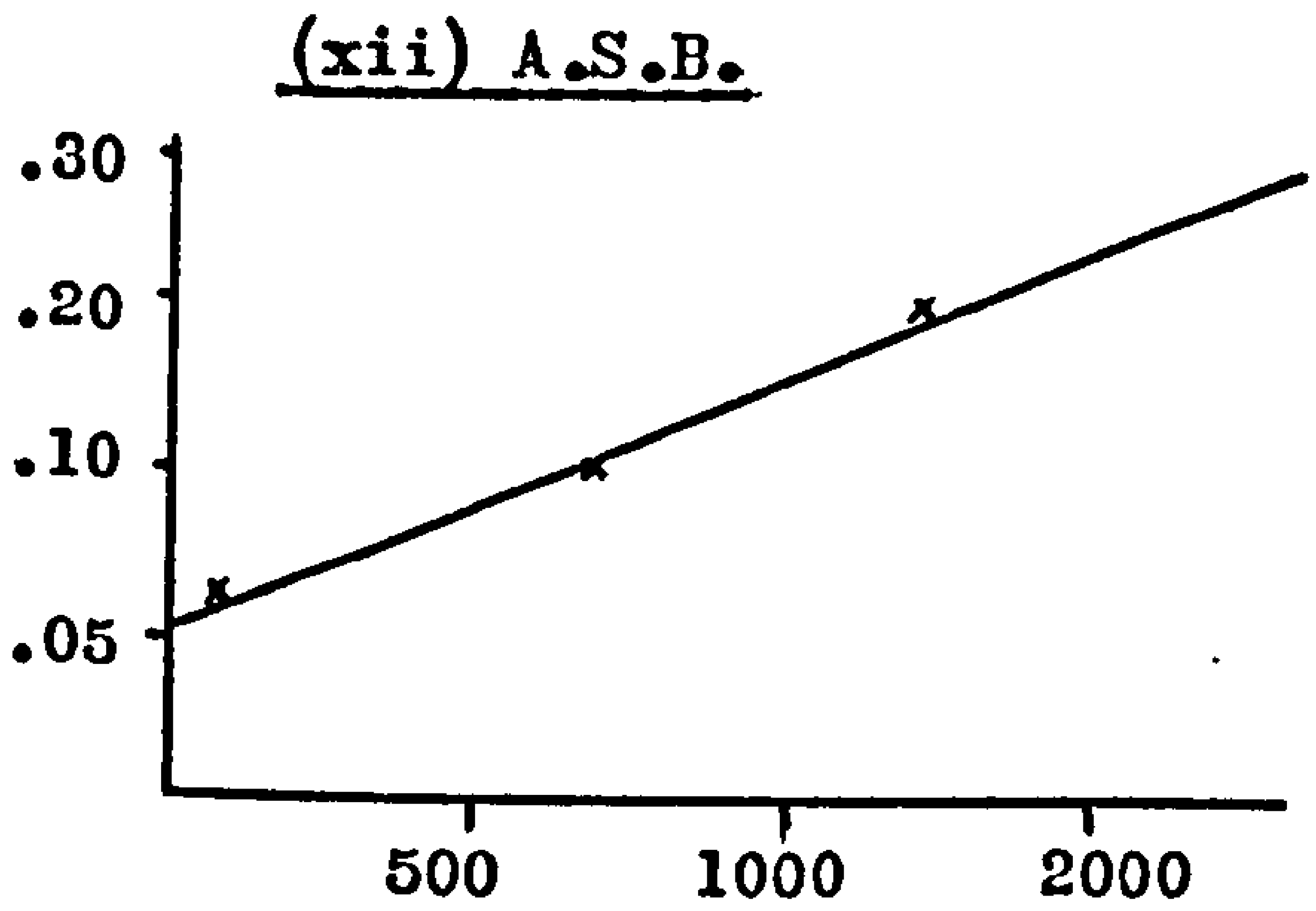


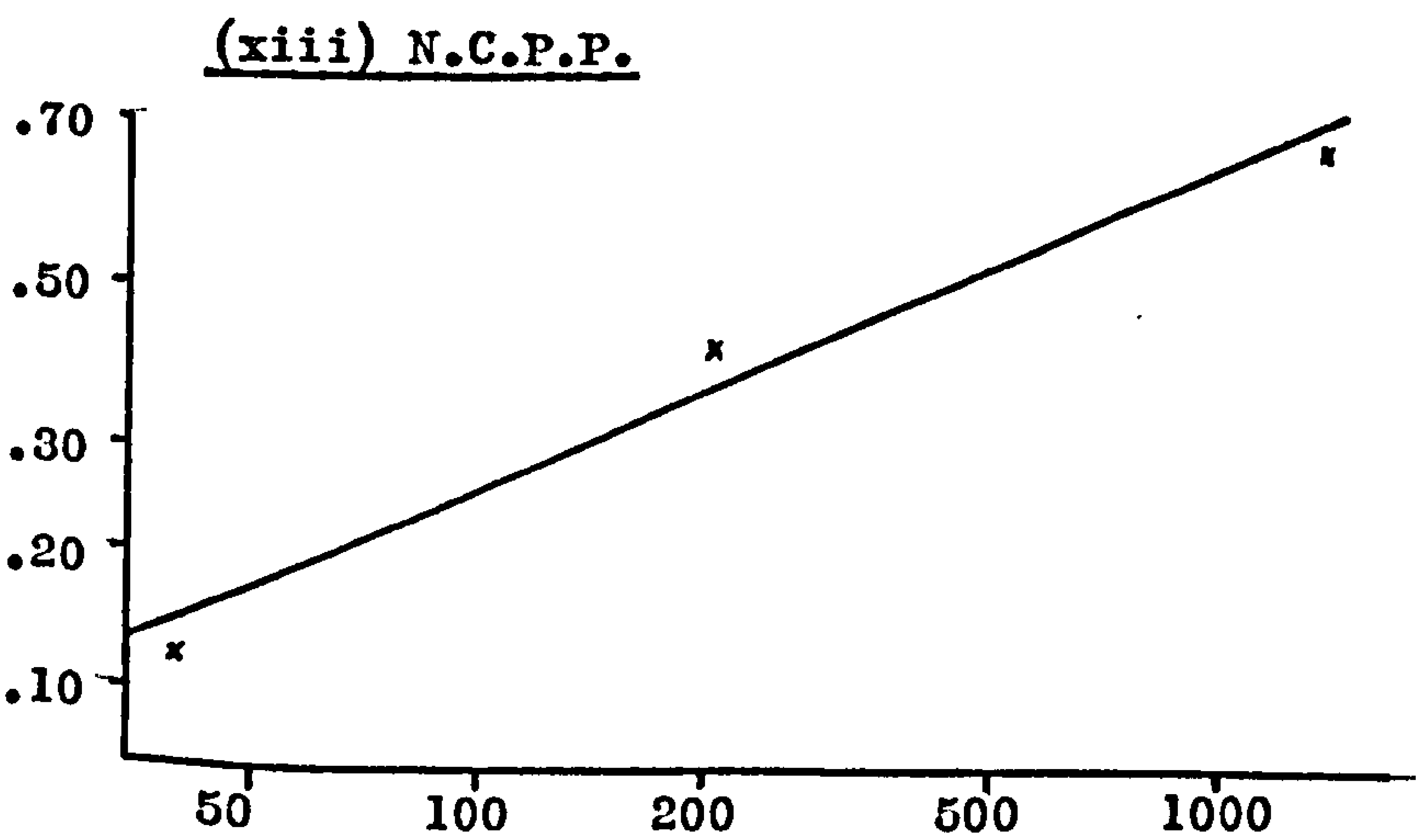

(viii) V.D.
.70
.50
.30
.20
.10
500
1000
2000
5000
10000
20000


(ix) V.M.
.90
.80
.70
.50
.30
.20
200
500
1000
2000
5000


(x) E.H.
.80
.70
.50
.30
.20
.10
.05
.01
50
100
200
500
1000


(xi) A.D.H.
.50
.40
.20
.10
.05
.01
50
100
200
500
1000

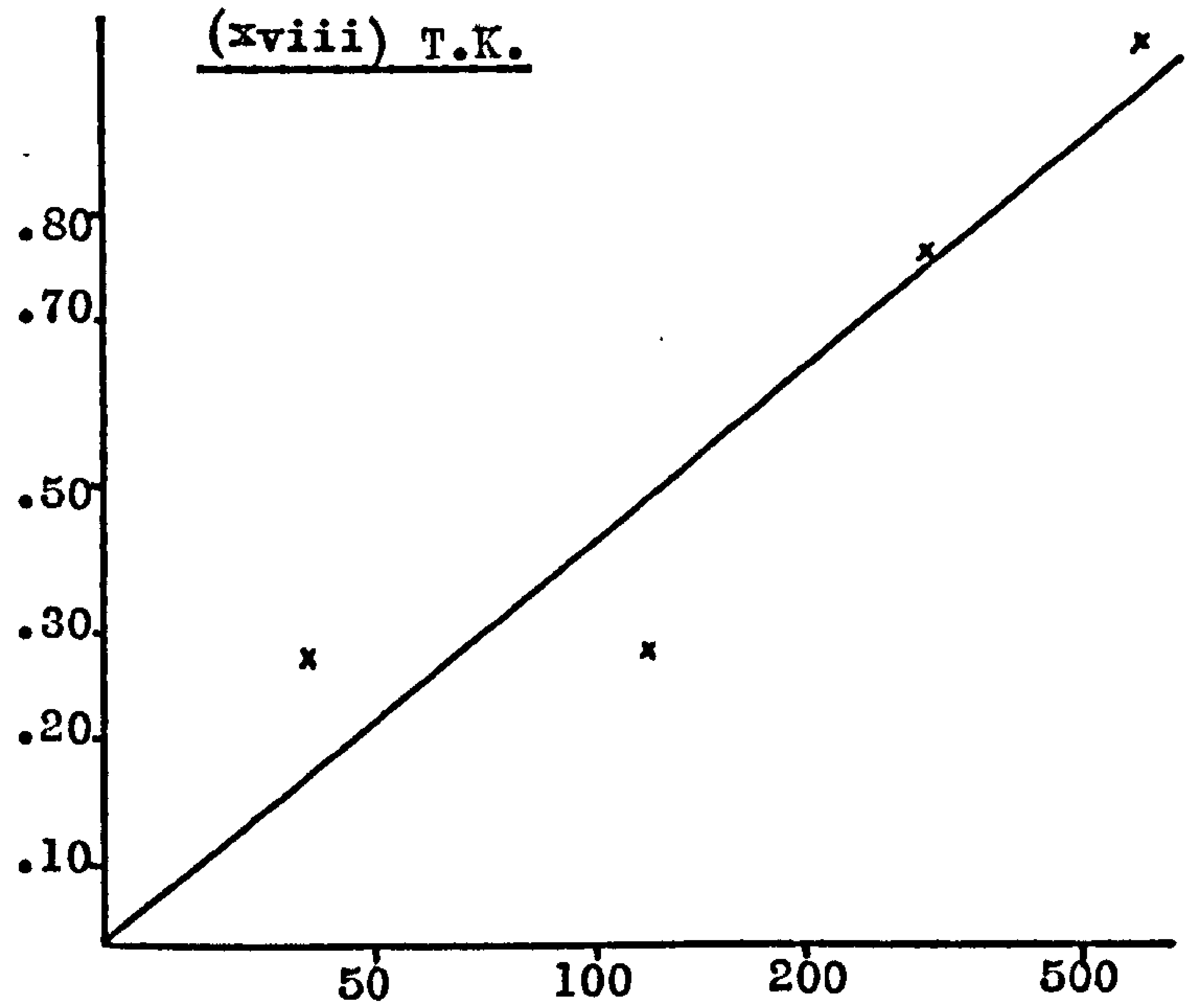

(xii) A.S.B.
.30
.20
.10
.05
500
1000
2000

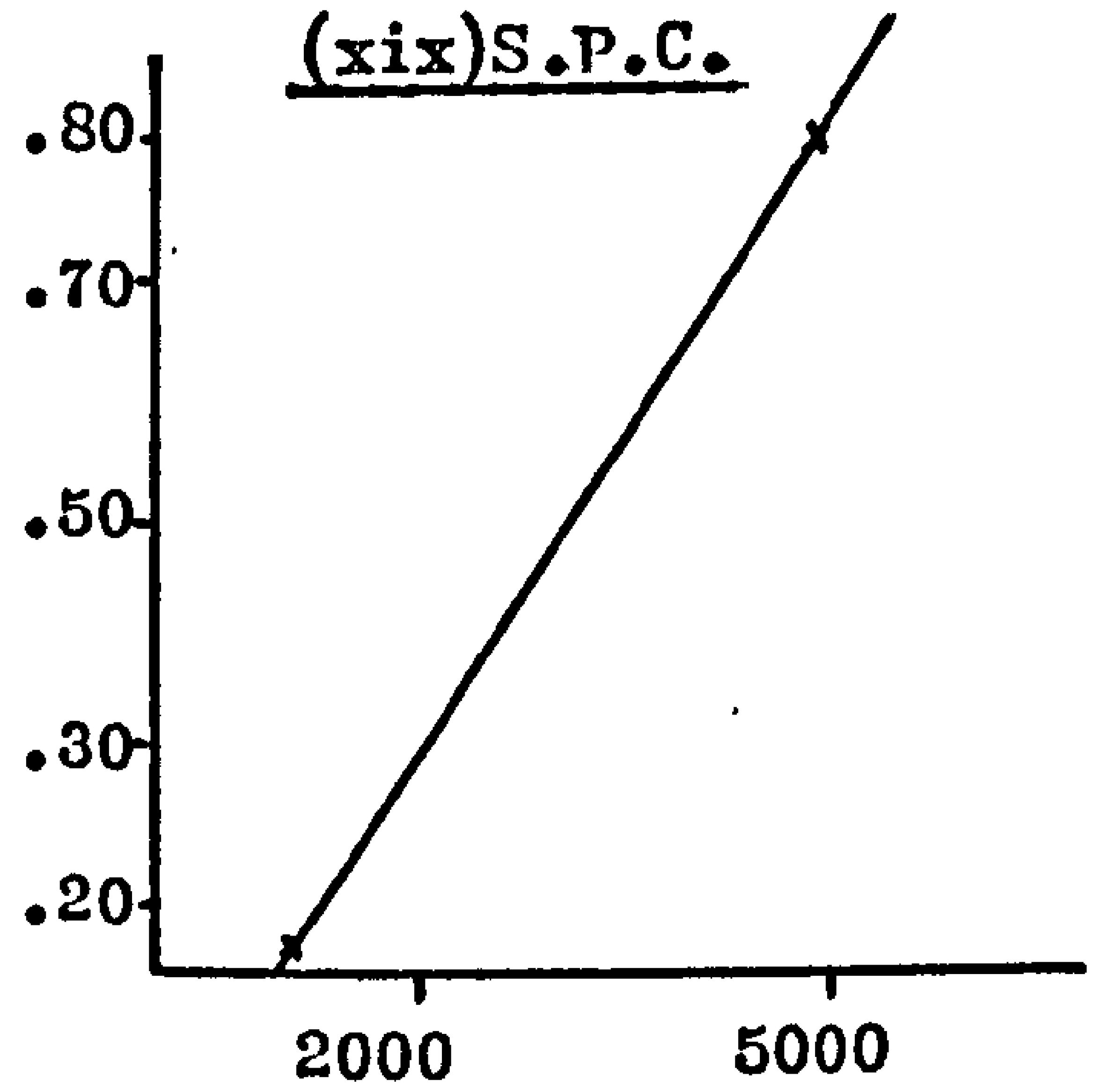

(xiii) N.C.P.P.
.70
.50
.30
.20
.10
50
100
200
500
1000

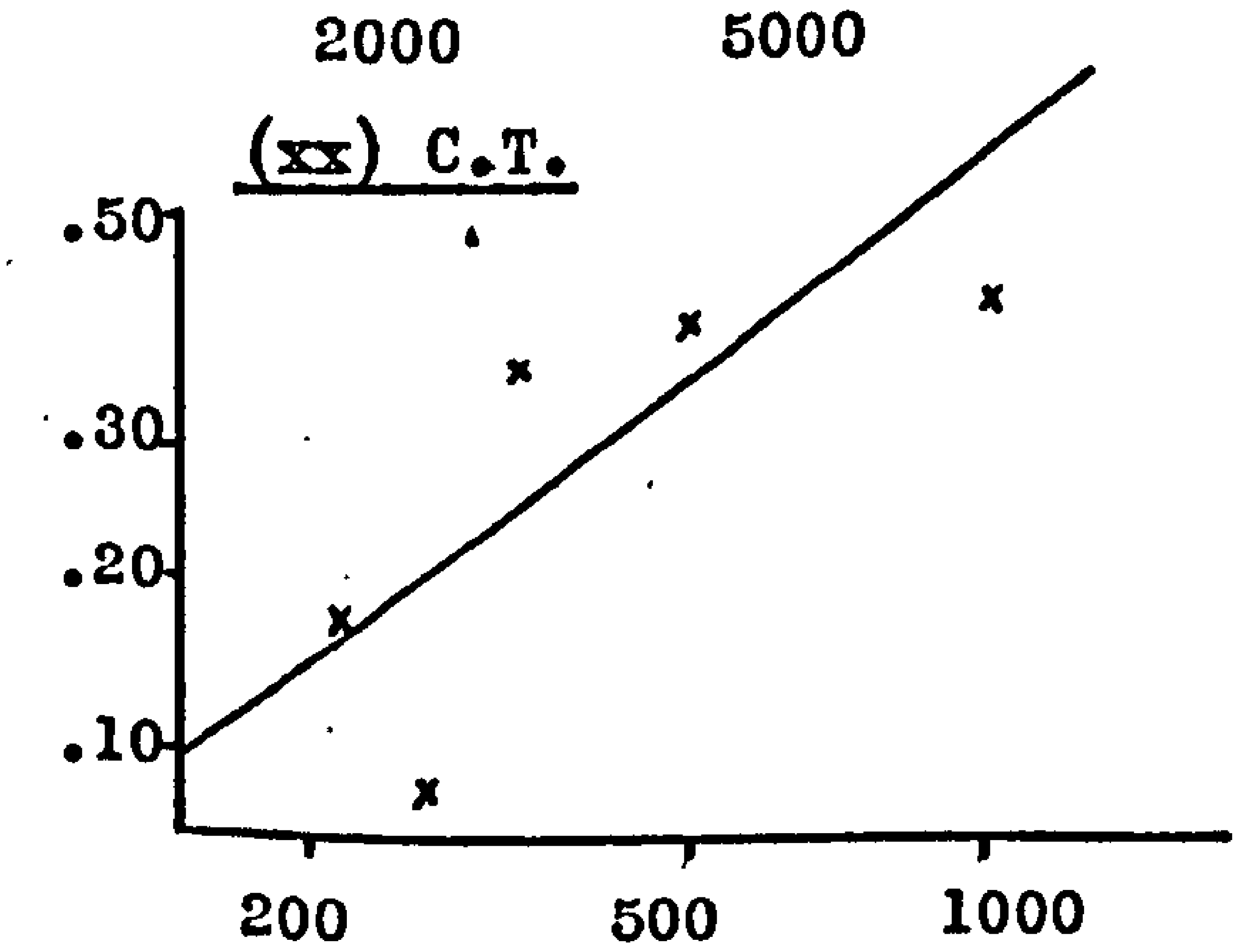
Ch.7.15.

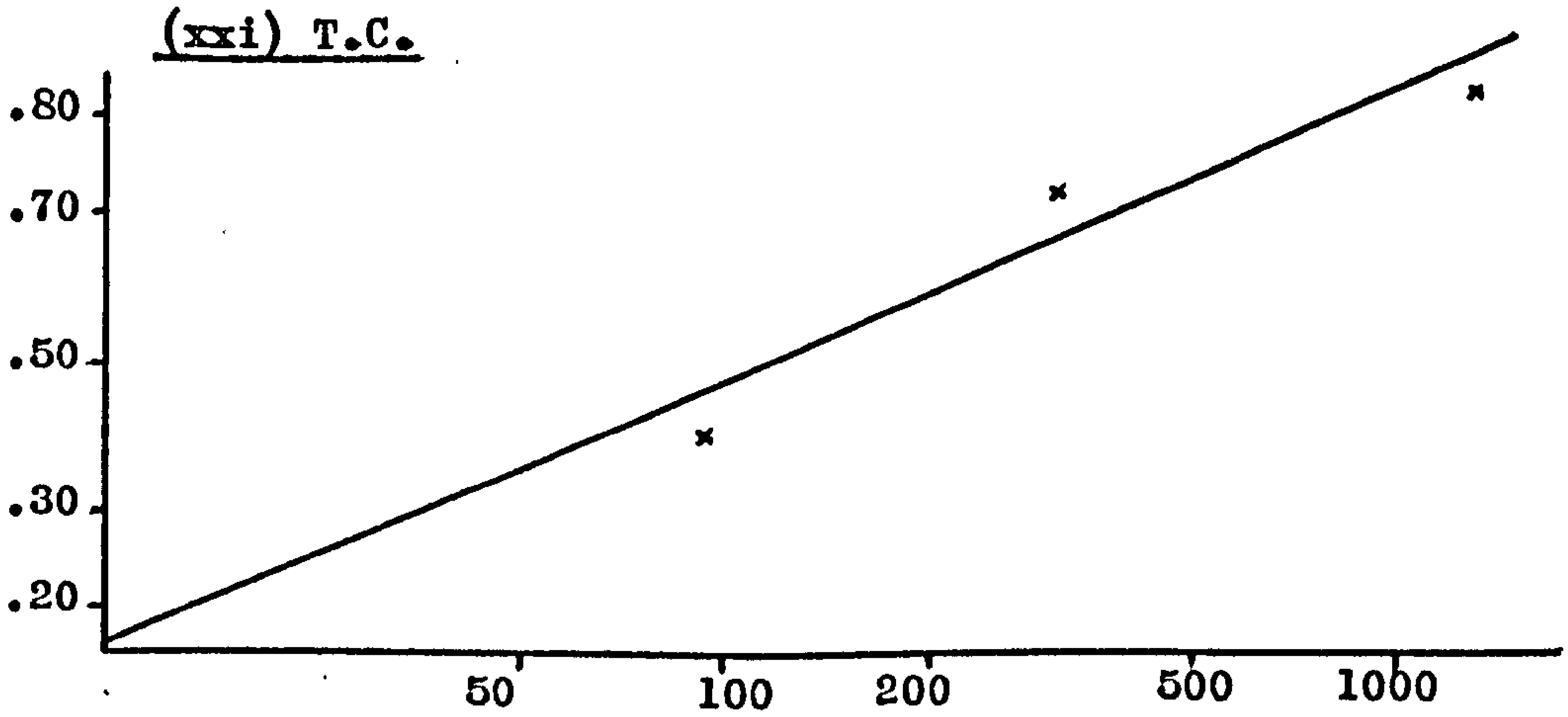

(xiv) NC.TURN.
.90
.80
.70
.50
.30
.20
.10
100
200
500
1000
2000
(xv) NC.TB.
.90
.80
.70
.50
500
1000
2000
5000
(xvi) G.A.
.90
.80
.70
.50
10
50
100
200
(xvii) S.L.
.90
.80
.70
.50
.30
.20
100
200
500
1000
2000

Ch.7.16.


(xviii) T.K.
.80
.70
.50
.30
.20
.10
50
100
200
500


(xix)S.P.C.
.80
.70
.50
.30
.20
2000
5000


(xx) C.T.
.50
.30
.20
.10
200
500
1000


(xxi) T.C.
.80
.70
.50
.30
.20
50
100
200
500
1000

Chapter 8. Cross - industry empirical analysis.

As the final stage in the empirical analysis, this chapter attempts an explanation of the causes for inter industry/innovation differences in the speed of diffusion. In doing so, it employs the estimated parameters of the diffusion growth curves and Quasi-Engel curves reported in previous chapters.

In section 1, for both cumulative normal and lognormal diffusion curves, measures are derived for 'the speed of diffusion' which can be computed from the parameters estimated in chapter 6. In section 2, these measures are shown to be somewhat complicated amalgams of the structural parameters of the theoretical model: α_j, β_j, Ψ_j, σ_j^2, μ_{sj} and σ_{sj}^2. As the model provides predictions as to the causes for inter industry/innovation differences in these parameters, the implications for differences in the speed of diffusion follow quite easily. There are, however, three obstacles to be overcome before econometric testing of these implications is possible. These are discussed, in turn, in section 3. First, the mathematical connection between the speed of diffusion and the structural parameters is sufficiently complicated to rule out direct estimation of the relationship between the speed of diffusion and its industry and innovation level determinants. Instead, it is necessary to break down the speed of diffusion into a number of constituent parts which may be more easily explained, independently, using straightforward regression techniques. This amounts to a transformation of the dependent variables using the cross-section estimates provided by the Quasi-Engel curves and empirical firm size distributions. Second, precise functional forms for the relationships between the structural parameters and their industry/ innovation level determinants must be assumed arbitrarily in the absence of any definite suggestions emanating from the model. Third, a number of the explanatory variables are somewhat intangible, or at least difficult to quantify, and suitable proxies must be decided upon. (Appendix 6 discusses, at some length, the problems of measuring the level of competition.) These obstacles resolved, four estimating equations (two for each growth curve) emerge to form the basis of the empirical analysis of this chapter. Sections 4 and 5 include the results of testing these equations for those innovations having cumulative normal and cumulative lognormal diffusion curves respectively. Section 6 reports two yardstick tests for these results.

1. The speed of diffusion.

It was shown in chapter 2 that given a logistic diffusion curve,

$$\left(\frac{m}{n}\right)_t = \left(1 + e^{-(\alpha + \beta t)}\right)^{-1} \qquad (8.1.1)$$

β may be usefully represented as the speed or rate of diffusion. It, alone, defines the time lapse between diffusion reaching two levels x_1 and x_2 where $1 > x_2 > x_1 > 0$. For example, the time span between diffusion attaining 20% and 80% is given by $\frac{\log 16}{\beta}$.[1]

Given, however, that the model predicts cumulative normal or lognormal diffusion, is there an equivalent parameter for these two curves which provides the same function as β in the logistic form?

The estimating forms used in chapter 6 were:

$$\bar{z}_t = a + b\log t + cC_t \qquad (8.1.2)$$

and

$$\bar{z}_t = a + bt + cC_t \qquad (8.1.3)$$

Using the above definition of the speed of diffusion (SD),[2]

$$SD = t_2 - t_1$$

where t_2 and t_1 denote the number of years elapsing before diffusion ($\bar{Q}_t$ as defined in chapter 6[3]) reaches $\bar{Q}_2 = x_2$ and $\bar{Q}_1 = x_1$ where $1 > x_2 > x_1 > 0$.

If $\bar{z}_1$ and $\bar{z}_2$ are normits of $\bar{Q}_1$ and $\bar{Q}_2$, then from (8.1.2):

$$\log t_1 = 1/b\,(\bar{z}_1 - a - cC_1) \qquad (8.1.4)$$

$$\log t_2 = 1/b\,(\bar{z}_2 - a - cC_2) \qquad (8.1.5)$$

As c was found to be not significantly different from zero for 20 of the 22 fitted cumulative lognormals, let $c = 0$,[4] in which case the expression:

1. page 2.4, chapter 2, section 1.
2. Thus SD is an inverse measure of the speed of diffusion.
3. Q_t is exactly equivalent to $(^m/_n)_t$
4. Alternatively, $C_1 = C_2 = 0$ has the same effect. Essentially, then, c only affects the speed of diffusion if capacity usage deviates in 1 and 2 from its level in year 0: Co has been defined as equal to 0.

$$SD_{LN} = t_2 - t_1 \qquad (8.1.6.)$$

reduces to

$$SD_{LN} = e^{-\frac{a}{b}} \left(e^{\frac{\bar{z}_2}{b}} - e^{\frac{\bar{z}_1}{b}} \right) \qquad (8.1.7)$$

For fixed values of $\bar{z}_1$ and $\bar{z}_2$, then, the speed of diffusion is determined by b and - a/b. Indeed, in the special case of $x_1 = 0$, and $x_2 = .50$ ($\bar{z}_1 = -\infty$ and $\bar{z}_2 = 0$).

$$SD_{LN} = e^{-a/b} .1. \qquad (8.1.8)$$

that is, $-a/_b$ uniquely determines the speed.

The picture is simpler for the cumulative normal, from (8.1.3)

$$SD_N = \frac{1}{b} (\bar{z}_2 - \bar{z}_1) - \frac{c}{b} (C_2 - C_1) \qquad (8.1.9)$$

Assuming that $c = 0$ or $C_2 = C_1$[1], this reduces, simply, to

$$SD_N = \frac{1}{b} (\bar{z}_2 - \bar{z}_1) \qquad (8.1.10)$$

and for fixed x_1 and x_2, the speed is determined, uniquely, by b.

As an example, if $x_1 = 20\%$ and $x_2 = 80\%$,

$$SD_N = \frac{1.68}{b}$$

Consequently, b will be used in the ensuing analysis as a measure of the speed of diffusion in the cumulative normal case. For the cumulative lognormal, -a/b is probably the more appropriate measure, as for many values of x_1, and x_2, b on its own, will have only a marginal influence on SD[2].

2. The implications for cross-industry explanations.

On the basis of these two measures, inter-industry/innovation differences in the speed of diffusion may be analysed as follows.

From (6.1.3), for the innovation in the jth industry:

$$a_j = \frac{(\log \alpha_j + \beta_j \mu_{sj})}{(\sigma_j^2 + \beta_j^2 \sigma_{sj}^2)^{\frac{1}{2}}} \qquad (8.2.1)$$

$$b_j = \frac{\Psi_j}{(\sigma_j^2 + \beta_j^2 \sigma_{sj}^2)^{\frac{1}{2}}} \qquad (8.2.2)$$

1. In this case, c only affects diffusion speed if capacity usage differs between the two years considered. In either case, however, it is surely correct to normalise for the effects of capacity usage.
2. For values of x_2 close to .50 and x_1 close to 0, the term in brackets in (8.1.7) tends to unity regardless of the value of b as $\bar{z}_2 \simeq 0$ and $\bar{z}_1 \simeq -\infty$ and thus, $SD_{LN} \simeq e^{-a/b}$.

and, therefore, $$-\frac{a_j}{b_j} = -\frac{(\log\alpha_j + \beta_j\mu_{sj})}{\Psi_j}. \qquad (8.2.3)$$

Thus, the speed of diffusion will vary between innovations depending on the values of the six structural parameters: α_j, β_j, Ψ_j, σ_j^2, μ_{sj} and σ_{sj}^2. These, in turn, are determined in terms of the model[1] by:

$$\alpha_j = g_1(\Pi_{oj}, CC_j, N_j, \Pi_j, TI_j, K_j, \bar{C}_j, VC_j) \qquad (8.2.4)$$

$$\beta_j = g_2(CC_j, TI_j, K_j) \qquad (8.2.5)$$

$$\sigma_j^2 = g_3(CC_j, GC_j, LI_j, TI_j, K_j, SC_j) \qquad (8.2.6)$$

$$\Psi_j = g_4(TI_j, CS_j, CC_j, N_j, \Pi_j, \bar{C}_j, K_j) \qquad (8.2.7)$$

μ_{sj} and σ_{sj}^2 are, of course, exogeneous to the system, being the parameters of the jth industry's size distribution.

A cross-section explanation of the estimated $\hat{b}_j$ and $\widehat{a_j/b_j}$ should be conceptually possible, therefore, given observations for all j of μ_{sj}, σ_{sj}^2 and the independent variables included in functions g_1 to g_4.

3. The estimating equations.

Substituting equations (8.2.4) - (8.2.7) into (8.2.2), then, the speed of diffusion will vary across industries having cumulative normal growth curves according to:

$$b_j = \frac{g_4(TI_j, CS_j, CC_j, N_j, \Pi_j, \bar{C}_j, K_j)}{\left\{g_3(CC_j, GC_j, LI_j, TI_j, K_j, SC_j) + [g_2(CC_j, TI_j, K_j)]^2\sigma_{sj}^2\right\}^{\frac{1}{2}}} \qquad (8.3.1)$$

and, for those industries having cumulative lognormal diffusion growth curves, according to:

$$-\left(\frac{a}{b}\right)_j = \frac{-\left\{\log g_1(\Pi_{oj}; CC_j; N_j; \Pi_j; TI_j; K_j; \bar{C}_j; VC_j) + g_2(CC_j; TI_j; K_j)\mu_{sj}\right\}}{g_4(TI_j; CS_j; CC_j; N_j; \Pi_j; \bar{C}_j; K_j)} \qquad (8.3.2)$$

Before these two relationships can be tested against the data, the precise form of functions g_1 to g_4 must be specified and empirical measures must be derived for the 12 different explanatory variables which appear in those functions. However, a potentially more troublesome problem arises

1. That is equations (5.6.1) - (5.6.4) of chapter 5.

concerning the method of estimation. As they stand, neither of these relationships may be estimated using standard linear regression techniques no matter how g_1 to g_4 are specified; there is no transformation of the right hand sides of (8.3.1) and (8.3.2) which could yield linear relationships. One way round this problem would be to ignore the precise mathematical forms of (8.2.2) and (8.2.3) and, instead, to assume general functions G_1 and G_2:

$$b_j = G_1(\Psi_j; \sigma_j^2; \beta_j; \sigma_{sj}^2) =$$

$$G_3\,(TI_j; CS_j; CC_j; N_j; \Pi_j; \bar{C}_j; K_j; GC_j; LI_j; SC_j; \sigma_{sj}^2) \quad (8.3.3)$$

and $$-{}^{a_j}/b_j = G_2\,(\alpha_j; \beta_j; \mu_{sj}; \Psi_j)$$

$$= G_4(\Pi_{oj}; CC_j; N_j; \Pi_j; TI_j; K_j; \bar{C}_j; VC_j; \mu_{sj}; CS_j) \quad (8.3.4)$$

Then G_3 and G_4 might be approximated by some simple functional form - say linear in logs - which might be estimated using ordinary least squares.

This possible approach is rejected, however, as it would largely reduce the model to an irrelevance in this area of the empirics. After all, equations (8.3.3) and (8.3.4) could be produced by purely ad-hoc reasoning: all of the independent variables, with perhaps the exception of μ_{sj} and σ_{sj}^2, might be suggested as determinants of the speed of diffusion on a purely intuitive level.

Fortunately, a more acceptable alternative is possible which uses explicitly the mathematical specifications of (8.2.2) and (8.2.3) and which is therefore much more within the spirit of the model. Basically, it amounts to a transformation of the dependent variables b_j and $-(a_j / b_j)$ so as to partly remove the non-linearity of (8.3.1) and (8.3.2).

3(i) Transformation of the dependent variables.

(8.2.2) may be re-written as:

$$b_j = (\Psi/\beta)_j \, (\sigma_j^2/\beta_j^2 + \sigma_{sj}^2)^{-\frac{1}{2}} \qquad (8.3.5)$$

and thus $$b_j\,({}^{\sigma_j^2}/_{\beta_j^2} + \sigma_{sj}^2)^{\frac{1}{2}} = (\Psi/\beta)_j \qquad (8.3.6)$$

For each industry, estimates are available for $(\sigma/\beta)_j^2$ and σ_{sj}^2 : the former is the variance of the Quasi-Engel curve, estimated in the previous chapter, and the latter is the variance of the logarithim of firms size in the jth industry, which is estimated in Appendix 5. Thus, it is possible, by combining the time series estimates for $(\hat{b}_j)$ and cross-section estimates $(\widehat{\sigma/\beta})_j$ and $(\hat{\sigma}_{sj}^2)$, to compute, for each industry, estimates of $(\widehat{\Psi/\beta})_j$, as in table (8.3.1). As this new dependent variable is determined across industries as the ratio of two general functions, g_4 and g_2, the only arbitrary element remaining in the ensuing estimating equation will be in the choice of specific forms for these functions. In addition, of course, the estimates of $(\widehat{\sigma/\beta})_j$ used in (8.3.6) to generate $(\widehat{\Psi/\beta})_j$ must themselves be explained across industries as $(\widehat{\Psi/\beta})_j$ is only one constituent part of $\hat{b}_j$.

The transformation for the cumulative lognormal speed of diffusion follows much the same lines. (8.2.3) may be re-written as:

$$-\,({}^{a}/_{b})_j = -(\frac{\log\alpha}{\Psi})_j - (\beta/\Psi)_j\,\mu_{sj} \qquad (8.3.7)$$

and thus,

$$-\,({}^{a}/_{b})_j + (\beta/\Psi)_j\,\mu_{sj} = -\,(\frac{\log\alpha}{\Psi})_j \qquad (8.3.8)$$

For each industry, estimates are available for $(\beta/\Psi)_j$ and μ_{sj}: the former may be computed from (8.3.6) (although the $\hat{b}_j$ used will be those generated by the cumulative lognormal curves in this case,) and the latter is the mean of the log. of firm size in the jth industry, which is estimated in Appendix 5. So, again, a combination of time series estimates $(-\,(\widehat{a/b})_j$ and $\hat{b}_j)$ and cross-section estimates $((\widehat{\sigma/\beta})_j$ and $\hat{\mu}_{sj})$ produces estimates of the transformed dependent variables as in table (8.3.1). As $-(\widehat{\log\alpha/\Psi})_j$ constitutes only one part of $-(\widehat{a/b})_j$, an empirical explanation of $(\widehat{\beta/\Psi})_j$ is also necessary, of course.

Table 8.3.1. The transformed dependent variables.

(a) Cumulative normal

	$\hat{b}_j$	$\widehat{(\sigma/\beta)}_j$	$(\widehat{\sigma_{sj}})^2$	$\widehat{(\Psi/\beta)}_j = \hat{b}_j\{(\sigma/\beta)^2 + \sigma^2_{sj}\}^{1/2}$
BOP	.148	0.96	1.145	.212
ATL	.179	2.81	.364	.514
TK	.093	1.21	2.756	.192
VM	.190	1.65	1.742	.401
PCBC	.142	2.62	2.941	.446
WSB	.076	2.90	2.941	.261
F	.310	1.80	2.941	.772
GA	.192	3.35	1.531	.685
TC	.046	2.38	3.842	.142
NCPP	.131	2.25	3.063	.373
CT	.205	1.34	.608	.319
NCTB	.161	2.23	1.960	.454
VD	.150	3.12	1.501	.503
SL	.116	1.69	2.890	.278
SPC	.113	.64	.701	.120

(b) Cumulative lognormal

	$-\widehat{(a_j/b_j)}$	$\hat{b}_j$	$\widehat{(\beta/\Psi)}_j$	$\hat{\mu}_{sj}$	$-\widehat{(\log\alpha/\Psi)}_j = -\widehat{(a/b)}_j + \widehat{(\beta/\Psi)}_j\mu_{sj}$
BOP	2.472	.752	.927	8.68	10.518
ATL	1.677	.882	.395	10.16	5.690
CC	3.131	.742	.338	8.45	5.987
PCBC	5.601	.424	.751	5.15	9.469
SP	2.731	1.494	.268	5.15	4.111
ASB	6.395	.339	1.069	4.67	11.387
F	2.149	2.152	.187	5.15	3.112
SF	2.122	1.268	.354	5.15	3.945
GA	1.742	.766	.366	3.10	2.877
ADH	5.048	.311	1.869	4.67	13.776
TC	6.496	.238	1.364	4.07	12.047
NCPP	2.621	.509	.689	5.23	6.224
CT	2.800	1.328	.485	5.82	5.623

Notes: $\hat{b}_j$ for the cumulative normal are taken from table 6.4.1. or table 6.6.1. The estimates from the latter are only preferred if they occur in regressions in which C_t is a significant variable. Otherwise, and therefore in the majority of cases, the estimates from table 6.4.1. are preferred. Similarly $\widehat{(-a/b)}_j$ for the cumulative lognormal are from tables 6.5.1 and 6.6.1. $\widehat{(\sigma/\beta)}_j$ are from table 7.2.1. $\hat{\mu}_{sj}$ and $\hat{\sigma}^2_{sj}$ are from Appendix 5.

To summarise, for each innovation having a cumulative normal diffusion curve, the speed of diffusion, $\hat{b}_j$, is split into its three constituent parts: $(\widehat{\Psi/\beta})_j$, $(\widehat{\sigma/\beta})_j$ and $\hat{\sigma}_{sj}$ and the first two are to be econometrically explained separately. Likewise, where the diffusion curve is cumulative lognormal, the speed of diffusion, $-(\widehat{a/b})_j$, is split into its three constituent parts, $-(\widehat{\log\alpha/\Psi})_j$, $(\widehat{\beta/\Psi})_j$ and $\hat{\mu}_{sj}$ and, again, the first two are to be econometrically explained separately. It will be possible, in both cases at a later stage, to 're-assemble' the speed of diffusion and to provide an overall explanation of it.

3(ii) Specific functional forms:

The potential estimating equations may be expressed as follows:

$$(\widehat{\Psi/\beta})_j = \frac{g_4\ (TI_j;\ CS_j;\ CC_j;\ N_j;\ \Pi_j;\ \bar{C}_j;\ K_j)}{g_2\ \ (CC_j;\ \ TI_j;\ \ K_j)} = F_1(TI_j;CS_j;CC_j;\ N_j;\Pi_j;\bar{C}_j;K_j) \quad (8.3.9)$$

for both cumulative normal and lognormal curves,[1]

$$(\widehat{\sigma/\beta})_j = \frac{[g_3(CC_j;GC_j;LI_j;TI_j;\ K_j;SC_j)]^{\frac{1}{2}}}{g_2(CC_j;TI_j;\ K_j)} = F_2(CC_j;GC_j;LI_j;TI_j;K_j;SC_j) \quad (8.3.10)$$

for cumulative normal curves,

$$-(\frac{\log\alpha}{\Psi})_j = \frac{-\log[g_1(\Pi_{oj};CC_j;N_j;\ \Pi_j;TI_j;K_j;\ \bar{C}_j;VC_j)]}{g_4(TI_j;CS_j;CC_j;\ N_j;\ \Pi_j;\ \bar{C}_j;\ K_j)}$$

$$= F_3\ (\Pi_{oj};CC_j;\ N_j;\ \Pi_j;TI_j;\ K_j;\ \bar{C}_j;VC_j;CS_j) \quad (8.3.11)$$

for cumulative lognormal curves.

Up to this point the precise mathematical specifications of g_1 to g_4 have been left open: there is nothing in the model to suggest any specific form for these functions. Therefore whatever formulation is selected for F_1, F_2 and F_3 must be somewhat arbitrary. Clearly, for linear regression techniques to be applicable, a fairly simple form is preferable.

1. The equation must be estimated separately for the two sets of estimates as Ψ_j does not have the same meaning in the different curves. In the cumulative lognormal, $\Psi_j = \frac{dz_j}{dt}\frac{1}{\bar{z}_{jt}t}$ and in the cumulative normal, $\Psi_j = \frac{dz_j}{dt}\frac{1}{\bar{z}_{jt}}$.

Probably the most appropriate choice for F_1, F_2 and F_3 is the familiar Cobb-Douglas or linear-in-logs specification. This may be rationalised on two grounds: (i) because F_1, F_2 and F_3 are each ratios of two other unspecified functions, the partial derivatives with respect to any of the independent variables must depend on the values of the other independent variables [1] - this suggests some sort of multiplicative form as an approximation, (ii) of the conventional multiplicative functional forms, the linear-in-logs is perhaps, the most flexible in terms of the signs of the second order derivatives.[2]

As an example then, F_1 will be approximated by:

$$F_1 = A_o TI_j^{A_1} CS_j^{A_2} CC_j^{A_3} N_j^{A_4} \Pi_j^{A_5} \bar{C}_j^{A_6} K_j^{A_7}$$

It is readily acknowledged that this is, by definition, an almost arbitrary choice.

3(iii) Measurement of explanatory variables.

g_1, g_2, g_3 and g_4 together include 12 separate variables which need to be measured in order for estimation to be possible (plus σ^2_{sj} and μ_{sj} computed from the empirical size distributions in Appendix 5.) These 12 variables fall into 2 groups: industry-level characteristics (CC_j; N_j; $\bar{C}_j$; VC_j;GC_j;LI_j;SC_j; and CS_j) and innovation-level characteristics (K_j; Π_{oj}; Π_j and TI_j).

1. As an example, consider $F_1 = g_4/g_2$, $\partial F_1/\partial CC = \partial g4/\partial CC \cdot 1/g2 - g4/g2 \cdot \partial g2/\partial CC$ $= 1/g2(\partial g4/\partial CC - \partial g2/\partial CC \cdot g4/g2)$, even if g_4 and g_2 were linear functions and thus $\partial g4/\partial CC$ and $\partial g2/\partial CC$ both constant, the value of $\partial F1/\partial CC$ would still be determined by all of the other variables in g_2 and g_4. Approximating F_1 by $\prod_{k=1}^{7} X_k^{A_k}$(i.e. the Cobb-Douglas) where X_k refer to the 7 arguments in g_2 and g_4 (and $CC=X_3$), then $\partial F1/\partial X_3 = (\Pi X_k^{A_k})\, A_3/X_3$, that is, the partial derivative with respect to CC will be, similarly, dependent on the values of all other X_k.

2. Using the example of the previous footnote, it can be shown that the second order derivative, $\partial^2 F/\partial CC^2$, may have a different sign from $\partial F_1/\partial CC$, even if g_4 and g_2 are, themselves, linear functions. Whilst the linear-in-logs form does allow the second order partial derivative of any variable to have the opposite sign to the first order derivative, this is not true for the other two commonly used multiplicative forms: $F_1 = \Pi\, e^{A_k X_k}$ or $e^{F_1} = \Pi\, X_k^{A_k}$.

Before discussing how these may best be measured, three general considerations should be mentioned. First, given that there are at most 22 observations (and usually rather less) for the dependent variables, degrees of freedom are at a premium: 12 explanatory variables are clearly 'too many' in this context. Second, some of the variables are so similar (for instance, π_{oj} and π_j and K_j and TI_j) that there is a good case for amalgamating them into one empirically feasible measure (especially as π_{oj} and TI_j would be difficult to measure anyway.) Third, at least one variable, GC_j (geographical concentration), but also to a certain extent SC_j, requires far better quality data than is in fact available. As GC_j is considered to have probably only a minor role, it will not be included in the ensuing empirics.

In the light of these considerations the number of explanatory variables has been reduced (hopefully without much loss) to the seven empirical measures (plus μ_{sj} and σ^2_{sj}) described below. In each case, data problems and sources are mentioned alongside the connection between the empirical variable and its theoretical counterpart(s).

(a) Innovation-level variables.

<u>π_j : the typical pay-back associated with adoption of the innovation in industry j.</u>

π_j is used as an empirical measure of the typical profitability of the innovation, which appears as an argument in functions g_1 and g_4. In addition it is used as the best available measure of π_{oj} (the initial profitability of the innovation as claimed by its manufacturers) which also appears in g_1; unfortunately, reliable widespread information on π_{oj} is just not available. As the extent of post -invention improvements will vary from innovation to innovation, there will not be a direct proportional relationship between π_{oj} and π_j but, nevertheless, they should be fairly highly correlated across industries.

π_{ij} is measured, where possible, by the pay back achieved from adoption by an average sized firm at approximately midway in the time series estimation period, (i.e. typical, both across time and space.) Often this ideal measure is not possible and what seems to be the most typical estimate available is chosen. Appendix One discusses, for each innovation individually, the available estimates, whose sources are usually the manufacturers, consuming firms and scientific and trade journals. Payback is defined simply as the number of years required for the extra revenues generated by adoption to pay back the initial investment outlay; thus it is inversely related to the profitability of the innovation.

<u>K_j : the typical capital outlay required for adoption of the innovation in industry j.</u>

K_j is used as an empirical measure for two theoretical variables: K_j and TI_j (the technical complexity of the innovation), each of which appear in ε_1, ε_2, ε_3 and ε_4. Fairly certainly the more costly is an innovation, the more technically complex it is likely to be; an intuitive ranking of the sample innovations by their technical sophistication supports this hypothesis. At any event, alternative proxies for TI_j, such as capital intensity or the necessary skilled labour complement, would prove difficult to measure in practice.

K_j is measured, where possible, by the average capital outlay required for adoption where, again, 'average' is interpreted spatially and temporally. Whilst much more data is available for K_j than for π_j, this ideal average is not always possible and thus the most typical estimate is often used. The estimates and sources are given for innovations individually in Appendix One. The units of measurement are thousands of pounds. (Experiments were performed at a late stage with an alternative specification of 'relative' cost, in which K_j was 'normalised' by the average size of potentially adopting firms, measured by employees. The explanatory power of this alternative measure proved to be almost identical to that of K_j - absolute size.)

(b) Industry level variables.

LI_j : the labour intensity of the industry.

LI_j is used as an empirical measure of SC_j - the extent of the science base of industry j (which appears in g_3) and LI_j - the labour intensity of industry j (which also appears in g_3.) It is measured by the share of value added (net output) allocated to wages and salaries at the mid-point of the diffusion period; data is obtained from various Census of Production M.L.H. reports, usually 1958 or 1963.

It is used as an inverse proxy for SC_j, largely on the basis of a number of findings of Carter and Williams in the late 1950's. It is argued that when the wage share is high, mechanisation is likely to be low and thus the industry concerned is more likely to be characterised by a craft, rather than a science, base. Alternative proxies which might have been used, had sufficiently disaggregated data been available, are the proportion of managers or work force qualified as scientists or engineers and the proportion of total costs accounted for by Research and Development expenditures.

Because LI_j reflects two independent theoretical variables which have opposite effects,[1] the expected sign of $\partial g_3/\partial LI_j$ is unknown.

$\bar{C}_j$: the average level of capacity usage over the diffusion period in the industry.

$\bar{C}_j$ enters as an argument, in both functions g_1 and g_4, as representative of the typical state of demand and level of capacity usage respectively. It is measured, simply, by averaging C_t over the diffusion period; the sources of data are discussed in Appendix 3. (At a preliminary stage, alternative measures such as the annual growth rate of demand and the investment - sales - ratio were computed, but these were not used in the ensuing empirics due to a lack of adequate available data.)

1. See chapter 5, section 6(iii) for the expected influence of labour intensity in its own right.

VC_j : the variance in capacity usage in the industry over the diffusion period.

VC_j enters the analysis as a determinant of the attitudes to risk in g_1, in other words, it is a proxy for the state of uncertainty in industry demand. It is computed, for each industry, from the data for C_t, the sources for which are listed in Appendix 3. Given the method used for computing C_t, the implication of this variable is that an industry which grows at exactly x% per annum will face perfect certainty (VC = 0) as opposed to high uncertainty, reflected in high VC_j, for an industry with demand which fluctuates wildly around a basically static level.

There are, of course, other potential measures of demand uncertainty,[1] but these all require the collection of large amounts of extra (perhaps unavailable) data.

N_j : the number of potential consumers in the industry.

N_j is included explicitly in g_1 and g_4: it is argued that in industries with large N, information will be initially more sparse and will also improve at a slower pace. In addition, however, N_j may be thought of as representing at least one dimension of competition in the consuming industry (CC_j). It is conventional to argue that the larger is N, the more effective is competition, other things being equal. In Appendix 6, the important problem of how to measure CC_j is discussed at some length. It is suggested that there can be no unique measure, but rather that the three parameters of the firm size distribution (N_j, σ^2_{sj} and μ_{sj}) each represent different aspects of 'the level of competition'. Thus N_j is used as one measure of CC_j such that $\partial CC_j / \partial N_j > 0$. In which case, N_j is also present, implicitly, in g_1; g_2; g_3; g_4. As the diffusion data used excludes firms which enter or leave the industries during the diffusion period, there is no problem of <u>when</u> to measure N_j : estimates are taken from Appendix 5.

1. See, for example, D. Schwartzman, 'Uncertainty and the size of firm,' Economica, 1963.

NS_j : the number of firms supplying the innovation to the consuming industry.

NS_j is used as an empirical measure of CS_j - competition in the supplying industry - which appears in g_4. It is assumed that competition is more effective the more firms supply the innovation, thus $\frac{\partial CS}{\partial NS} > 0$. Clearly, this is an imperfect measure but more sophisticated alternatives are ruled out by the lack of widespread information on individual firms' market shares. In any event as NS_j is often very small,[1] measures such as 5 firm concentration ratios would not be fine enough to differentiate between most of the industries. Estimates of NS_j are available in Appendix 1.

σ^2_{sj}: the variance of the logarithim of firm size in industry j.

σ^2_{sj} does not appear explicitly in any of g_1, g_2, g_3 or g_4, but estimates are required in order to compute the dependent variables as defined in (8.3.6) and (8.3.8). Moreover, it is used to represent another dimension of competition in the consuming industry (CC_j); as explained in Appendix 6, the higher is σ^2_s, the lower should be the degree of effective competition, other things being equal. High σ^2_s will be associated, statistically, with industries exhibiting great inequality in firm size: where a handful of firms dominate, σ^2_{sj} will be high, for given N_j. Thus $\frac{\partial CC}{\partial \sigma^2_s} < 0$, and implicitly, therefore, σ^2_s is included in each of g_1, g_2, g_3 and g_4. Estimates are presented in Appendix 5 on the assumption, seemingly justified, that all of the sample size distributions may be approximated by the lognormal.

$e^{\mu_{sj}}$: the geometric mean firm size in industry j.

The argument here is exactly analagous to that for σ^2_s. Estimates of μ_{sj} are required to compute one of the dependent variables (see 8.3.8) and μ_{sj} may be seen as an implicit argument in g_1, g_2, g_3 and g_4, in that it represents a third dimension of competition in the consuming industry. In Appendix 6, it is argued that it represents certain barriers to entry such that $\partial CC / \partial e^{\mu_s} < 0$, that is, the higher is mean-firm size, the higher will be certain entry barriers and the less effective competition.

1. See section 5 of chapter 4.

Estimates of μ_{sj} are presented in Appendix 5; the only measurement problem is the choice of year in which to estimate - arbitrarily, the most recent year for which data is available is selected. Given the large differences in industry size, it is unlikely that the ranking of industries by μ_s will be sensitive to the year of observation.

On the basis of these empirical considerations, equations (8.2.4) - (8.2.7) are therefore approximated, for purposes of estimation by:

$$\alpha_j = g_1\,(\underset{-}{\Pi_j};\ \underset{-}{K_j};\ \underset{+}{\bar{C}_j};\ \underset{-}{VC_j};\ \underset{-}{N_j};\ \underset{+}{CC_j}) \qquad (8.3.12)$$

$$\beta_j = g_2\,(\underset{+}{K_j};\ \underset{+}{CC_j}) \qquad (8.3.13)$$

$$\sigma_j^2 = g_3\,(\underset{?}{LI_j};\ \underset{+}{K_j};\ \underset{?}{CC_j}) \qquad (8.3.14)$$

$$\Psi_j = g_4\,(\underset{+}{K_j};\ \underset{(-)}{\Pi_j};\ \underset{(+)}{\bar{C}_j};\ \underset{-}{N_j};\ \underset{?}{CC_j};\underset{?}{NS_j}) \qquad (8.3.15)$$

$$\Omega_j = g_5\,(\underset{?}{K_j};\ \underset{(-)}{\Pi_j};\ \underset{-}{N_j};\ \underset{-}{CC_j}) \qquad (8.3.16)$$

$$\text{where}\quad CC_j = g_6\,(\underset{-}{\mu_{sj}};\ \underset{-}{\sigma_{sj}^2};\ \underset{+}{N_j}) \qquad (8.3.17)$$

The expected signs of the partial derivatives are shown underneath each variable. Accordingly, the estimating forms for F_1, F_2 and F_3 are given by:

$$\overline{b_j\left(\sigma_j^2/\beta_j^2 + \sigma_{sj}^2\right)^{\frac{1}{2}}} = \widehat{(\Psi/\beta)}_j = A_0\,\underset{?}{K_j}^{A_1}\,\underset{?}{CC_j}^{A_2}\,\underset{(-)}{\Pi_j}^{A_3}\,\underset{(+)}{\bar{C}_j}^{A_4}\,\underset{-}{N_j}^{A_5}\,\underset{?}{NS_j}^{A_6} = F_{1j} \qquad (8.3.18)$$

$$\widehat{(\sigma/\beta)}_j = B_0\,\underset{?}{K_j}^{B_1}\,\underset{?}{CC_j}^{B_2}\,\underset{?}{LI_j}^{B_7} = F_{2j} \qquad (8.3.19)$$

$$\overline{(-a_j/b_j + (\beta/\Psi)\mu_{sj})} = -\left(\widehat{\frac{\log\alpha}{\Psi}}\right)_j = C_0\,\underset{?}{K_j}^{C_1}\,\underset{?}{CC_j}^{C_2}\,\underset{(+)}{\Pi_j}^{C_3}\,\underset{(-)}{\bar{C}_j}^{C_4}\,\underset{(+)}{N_j}^{C_5}\,\underset{?}{NS_j}^{C_6}\,\underset{+}{VC_j}^{C_8} = F_{3j} \qquad (8.3.20)$$

$$\text{where}\quad CC_j = D_0 N_j^{D_1}\, e^{D_2\sigma_{sj}^2}\, e^{D_3\mu_{sj}} \quad \text{(from Appendix 6)} \qquad (8.3.21)$$

It should be noted, in passing, that concrete expectations of signs are not always possible either because of alternative hypotheses or due to the possibility of cancelling. For instance, K_j is postulated as a positive influence on both σ^2 and β, therefore its overall influence on (σ/β) must be uncertain.

4. Results : the constituents of the cumulative normal speed of diffusion.

Table 8.4.1 presents the results of estimating equations (8.3.18) and (8.3.19) - each having been logarithimically transformed[1] - for two sub-samples of the innovations. Initially, the only innovations excluded from the full sample are those seven (CC, SP, ASB, SF, ADH, NCTURN and EH) for which the cumulative normal diffusion curve is an obvious mis-specification; VD, SL and SPC are each included even though autocorrelation existed when the cumulative normal was fitted,[2] on the grounds that in these cases the reasons for this may have been other than mis-specification of form.[3] However, in the second part of the table, results are reported for the sample excluding these three innovations.

Reducing the sample size in this way does reduce the number of degrees of freedom rather drastically: there are only 15 observations for the equations in part (a) and only 12 in part (b) of the table. Clearly, it is unrealistic to expect equations including up to seven variables to perform outstandingly under such circumstances: one outcome of this shortage in degrees of freedom is that a number of the explanatory variables are spuriously correlated.

When equation (8.3.18) is fitted to the first (larger) sample of observations, three variables, NS_j, K_j and μ_{sj} (as a measure of CC_j) quite clearly contribute nothing to the overall explanatory power.[4]

1. In all estimated equations, σ^2_{sj} and μ_{sj} are not logged as $\log CC_j = \log D_0 + D_1 \log N_j + D_2 \sigma^2_{sj} + D_3 \mu_{sj}$ from (8.3.21)

2. See table 6.4.1.

3. As was argued in section 7 of chapter 6, where autocorrelation is not caused by mis-specification or an omitted variable with a trend element, it is less likely to be accompanied by biassed estimates. Therefore, there are reasons for believing that $\hat{b}_j$ for these 4 innovations may not be biassed.

4. Not shown in the table.

Omitting these variables (equation 1 in the table) improves the fit as measured by the corrected R^2; of the four remaining explanatory variables, N_j and π_j have coefficients with the expected signs and significantly different from zero at the 95% level. Whilst σ^2_{sj} and $\bar{C}_j$ both remain insignificant, they cannot be confidently rejected as they are themselves highly positively correlated.[1] Therefore, to establish what happens in the absence of multi-collinearity, the equation is re-estimated excluding first σ^2_{sj} and then $\bar{C}_j$ (equations 2 and 3). As expected, the $\bar{R}^2$ remains virtually unchanged in both cases and each variable, in turn, has an increased t statistic, such that the weaker 90% test identifies their coefficients as significantly different from zero. However, in both cases the possibility of biassed estimates now arises. Equations 4 and 5 are estimated merely to establish that K_j and NS_j remain insignificant in even these revised equations. (When μ_{sj} is included alongside only N_j, π_j and σ^2_{sj}, it, too, adds nothing to the overall explanation.[2]) Finally, in equation 6 only π_j and N_j are included, and as can be seen they do provide the bulk of the explanation offered in equations 1 - 5.

Overall, these results are fairly encouraging: both π_j and N_j are robustly significant, regardless of which other variables are included in the estimating equation and $\bar{C}_j$ and σ^2_{sj} can certainly not be ruled out as explanatory variables : a more definite conclusion on these latter two might have been forthcoming given a larger sample size.

Unfortunately the results are not as impressive for $(\widehat{\sigma/\beta})_j$. When the full form of equation (8.3.19), including all three measures of CC_j, is estimated, only one variable, LI_j, has a significant coefficient and two others, μ_{sj} and K_j, have extremely low t statistics.[3] Omitting these two from the equation (equation 7) changes nothing, LI_j remains significant

1. With a coefficient of correlation of .7.

2. Not shown in the table.

3. Not shown in the table.

whilst the other two variables, N_j and σ^2_{sj}, still have low t statistics. Omitting each of these in turn (equations 8 and 9) does not improve the significance of the other, and it must be concluded that the fit achieved from the inclusion of only LI_j (equation 10) is probably the most satisfactory.

In part (b) of the table the results are reported of fitting the most interesting of the ten above equations to the smaller sample of innovations, excluding VD, SL and SPC. As such, this may be seen as a crude form of sensitivity analysis. In general, the results hold up well, even although there are now only 12 observations: R^2 usually improve although $\bar{R}^2$ sometimes decline, N_j and $\widehat{\Pi}_j$ are still significant determinants of $(\widehat{\Psi/\beta})$, and so is LI_j for $(\sigma/\beta)_j$. On the other hand, the t statistics for $\bar{C}_j$ and σ^2_{sj} decline sufficiently for them to be no longer quite significant at the 90% level, and although the t statistics for σ^2_{sj} and N_j improve slightly in equation 15, neither is significant at even the 90% level. Perhaps most important, however, is the fact that the coefficients of N_j, Π_j and LI_j seem fairly stable.

Overall then, three consistently significant determinants of the cumulative normal speed of diffusion have been established (N_j, Π_j and LI_j) and two other variables ($\bar{C}_j$ and σ^2_{sj}) cannot be entirely ruled out of the reckoning. For both of the constituent parts of the rate of diffusion there still remain large elements unexplained by the explaratory variables considered. This is not too surprising, given that they are, themselves, only estimates generated by the empirical work of the previous chapters; consequently they may contain large errors in measurement.

The economic implications of these results are investigated at some length in the following chapter.

Table 8.4.1. The cumulative normal speed of diffusion.

a) Including only those innovations for which the cumulative normal is not an obvious mis-specification of form. i.e. n = 15.

Dependent variable.	Const.	$\log N_j$	$\log \Pi_j$	$\log \bar{C}_j$	σ^2_{sj}	$\log K_j$	$\log NS_j$	$\log LI_j$	R^2	$\bar{R}^2$
$\log \widehat{(\Psi/\beta)}_j$										
1.	1.385^{++} (1.31)	-.437 (5.44)	-.323 (2.30)	6.387^{++} (1.46)	.108^{++} (.99)				.690	.577
2.	1.208 (2.78)	-.517 (5.38)	-.345 (2.36)	7.091^{+} (1.79)					.672	.599
3.	1.239 (2.21)	-.471 (4.84)	-.282 (2.31)		.339^{+} (1.90)				.671	.599
4.	1.091^{+} (1.81)	-.498 (4.05)	-.264^{+} (2.05)		.299^{++} (1.71)	-.031^{++} (0.59)			.673	.554
5.	1.239 (2.21)	-.511 (4.84)	-.243 (2.30)		.309^{+} (.182)		-.146 (0.41)		.685	.570
6.	0.799^{+} (2.05)	-.351 (4.86)	-.374 (3.48)						.614	.555
$\log \widehat{(\sigma/\beta)}_j$										
7.	.347^{++} (.62)	-.113^{++} (1.21)			.143^{++} (.79)			-1.083 (2.28)	.341	.176
8.	.185^{++} (.36)	-.076^{++} (0.95)						-1.234 (2.42)	.297	.139
9.	-.241^{++} (.71)				.030^{++} (.29)			-1.292 (2.45)	.267	.154
10.	-.203^{++} (.66)							-1.331 (2.68)	.264	.211

b) Including only those innovations in the cumulative normal acceptable set i.e. n = 12.

Estimated coefficients of independent variables.

Dependent variable.	Const.	$\log N_j$	$\log \Pi_j$	$\log \bar{C}_j$	σ^2_{sj}	$\log LI_j$	R^2	$\bar{R}^2$
$\log \widehat{(\Psi/\beta)}_j$								
11.	1.315^{++} (1.79)	-.433 (2.35)	-.311 (2.63)		.232^{++} (1.65)		.679	.559
12.	1.570^{+} (2.00)	-.533 (3.13)	-.332 (3.08)	8.311^{++} (1.73)			.689	.572

Table 8.4.1... continued ...

Dependent variable.	Const.	$\log N_j$	$\log \Pi_j$	$\log \bar{C}_j$	σ^2_{sj}	$\log LI_j$	R^2	$\bar{R}^2$
$\log \widehat{(\Psi/\beta)}_j$								
13.	.830 (0.93)	-.325 (2.47)	-.376 (3.25)				.573	.478
$\log \widehat{(\sigma/\beta)}_j$								
14.	-.039++ (.97)					-1.032 (2.34)	.352	.287
15.	.466++ (.86)	-.306++ (1.56)			.235++ (1.41)	-1.327 (2.31)	.430	.216

Notes: Figures in parentheses beneath estimated coefficients are t statistics.

+ denotes a coefficient not significantly different from zero at the 95% level.

++ denotes a coefficient not significantly different from zero at the 90% level.

5. Results: the constituents of the cumulative lognormal speed of diffusion.

Table 8.5.1 presents the results for equation (8.3.20) and, again, (8.3.18), but this time using the values of $\widehat{b}_j$ and $(\widehat{-a/b})_j$ derived from the cumulative lognormal diffusion curves. The sample comprises only those innovations for which the cumulative lognormal provided an acceptable fit in chapter 6.[1] Thus, 9 innovations are excluded from the full sample (TK, VD, VM, WSB, EH, NCTB, SL, SPC, NCTURN), as there were strong reasons for believing that the cumulative lognormal constituted a mis-specification of their diffusion curves.

As can be seen from the table, the results are very similar to those reported in the previous section. For equation (8.3.20.), four of the potential explanatory variables, K_j, NS_j, VC_j and μ_{sj}, appear to have no influence on $-(\frac{\log\alpha}{\Psi})_j$ when all eight variables are included in the same equation.[2] When they are omitted and the equation is re-estimated (equation 1), the corrected R^2 improves slightly and both N_j and Π_j have coefficients with the expected sign and significantly different from zero, but $\bar{C}_j$ and σ^2_{sj} remain insignificant, even at the 90% level. As already stated, however, these two variables are collinear. Not surprisingly, when σ^2_{sj} is then omitted, (equation 2), $\bar{C}_j$'s t statistic does improve, but not sufficiently to establish the significance of its coefficient. On the other hand, σ^2_{sj}'s t statistic hardly improves at all when C_j is omitted (equation 3). Equation 4, in which only Π_j and N_j are included, establishes again that these two variables account for nearly all the explanatory power of the previous three equations. In a series of (unreported) regressions, each of K_j, NS_j, VC_j and μ_{sj} have been added, in turn, to these two variables but none of them ever approaches significance.

The, by now familiar, pattern re-emerges when equation (8.3.18) is estimated, using the lognormal estimates of $(\widehat{\Psi/\beta})_j$: N_j and Π_j account for most of the explanation achieved, $\bar{C}_j$ and σ^2_{sj} record t statistics usually in excess of unity, but far short of the required critical level, and the remaining explanatory

1. See tables 6.7.2 and 6.5.1.

2. Not shown in the table.

Table 8.5.1. The cumulative lognormal speed of diffusion.

Based on observations for the cumulative lognormal set only, n = 13

Dependent variable.	Estimated coefficients of independent variables						
$\log\left\{-\left(\frac{\log\alpha}{\Psi}\right)_j\right\}$	Const	$\log N_j$	$\log \Pi_j$	$\log \bar{C}_j$	σ^2_{sj}	R^2	$\bar{R}^2$
1.	$-.031^{++}$	.245	.227	-8.178^{++}	$.102^{++}$	.829	.744
	(.60)	(2.52)	(2.06)	(1.49)	(.98)		
2.	$.234^{++}$	.232	$.183^{+}$	-8.980^{++}		.802	.736
	(.57)	(2.43)	(1.85)	(1.67)			
3.	$-.113^{++}$	.316	.317		$.125^{++}$	.766	.688
	(.21)	(3.48)	(3.16)		(1.12)		
4.	$.212^{++}$	.311	.273			.724	.669
	(.47)	(3.37)	(2.92)				
$\log \widehat{(\Psi/\beta)}_j$							
5.	1.739	-.409	$-.188^{+}$	6.964^{++}	$-.052^{++}$	.785	.678
	(.92)	(2.31)	(2.01)	(1.37)	(.75)		
6.	2.008^{++}	-.325	$-.192^{+}$	8.280^{++}		.748	.664
	(.62)	(2.29)	(1.92)	(1.43)			
7.	2.093^{++}	-.371	-.290		$-.095^{++}$	.722	.629
	(.35)	(3.17)	(3.02)		(1.01)		
8.	2.299^{++}	-.406	-.234			.702	.642
	(.64)	(3.24)	(2.54)				

Notes. Figures in parentheses below estimated coefficients are t statistics.

+ denotes a coefficient not significantly different from zero at the 95% level.

++ denotes a coefficient not significantly different from zero at the 90% level.

variables K_j, μ_{sj} and NS_j add nothing to the overall explanation.

Therefore, as in the previous section, results are <u>fairly</u> encouraging: corrected $\bar{R}^2$ exceed .5 and two variables prove to be significant determinants of both constituent parts of the cumulative lognormal speed of diffusion. However, there remains a substantial proportion of the variance in $-(\widehat{\log\alpha/\psi})_j$ and $(\widehat{\psi/\beta})_j$ that is unexplained by the explanatory variables suggested by the model.

6. Results: using untransformed time series parameters and the logistic speed of diffusion.

The results reported in the previous two sections, whilst encouraging, are somewhat inconclusive. Given the rather unusual form of the dependent variables,[1] it is not clear what would constitute an acceptable fit in these circumstances. To provide some sort of perspective, therefore, two alternative approaches have been pursued.

First, regressions are computed on the basis of the estimating equations suggested by (8.3.3) and (8.3.4). In this way, it is possible to assess the advantages (if any) to be gained from splitting up the speed of diffusion into its constituent parts. It will be recalled that these alternative estimating equations ignore the precise mathematical form of the relationship between the structural parameters and the speed of diffusion. In other words, the intermediate role of the structural parameters is by-passed, and the estimating equations relate the speed of diffusion directly to the explanatory variables. Assuming that a linear-in-logs specification adequately approximates both G_3 and G_4, (8.3.3) and (8.3.4) are estimated using the samples described, respectively, in the previous two sections.

When (8.3.3) is tested in full, none of the 11 explanatory variables has a significant coefficient at the 95% level, but then this is not surprising given that there are only 4 degrees of freedom. As a more appropriate test, the equation is re-computed, using only those variables which were significant

1. Namely, <u>estimates</u> of population parameters derived from time series regressions rather than observed data.

in section 4 plus σ^2_{sj} (equation 1). As can be seen, only N_j and π_j are significant at the 95% level and $\bar{C}_j$ at the 90% level; the corrected R^2 is substantially lower than that typically attained for $(\widehat{\Psi/\beta})_j$ but higher than for $(\widehat{\sigma/\beta})_j$. Perhaps most interesting, however, is the negative and insignificant coefficient on σ^2_{sj}. From the definition according to the model (equation 8.2.2), b_j should be inversely related to σ^2_{sj}, but on the other hand, there is some faint evidence, in section 4, that, acting as a measure of the degree of competition, it has a positive influence on $(\widehat{\Psi/\beta})_j$. Partly because these two effects may cancel each other out <u>on average</u>, but probably also due to the approximation of mathematical form used, neither of these influences are 'picked up' in equation 1. This tends to suggest that this alternative approach is perhaps not as sensitive, as well as having an inferior explanatory power. Similarly, the influence of LI_j (as a determinant of $(\widehat{\sigma/\beta})_j$) is not reflected significantly in equation 1.

The same sort of results appear when using this cruder approach to explain the lognormal speed of diffusion (equation 8.3.4). Again, when including all of the explanatory variables, none appear as significant.[1] Including only those variables which were shown to be significant in section 5, plus μ_{sj}, (equation 2), the corrected R^2 is lower than that achieved for $-(\widehat{\log\alpha/\Psi})_j$ and about the same as for $(\widehat{\Psi/\beta})_j$, but rather surprisingly, π_j is only significant at the 90% level. Interestingly, much of the explanatory power of the equation is due to the inclusion of μ_{sj}. This variable has proved to be totally insignificant as a determinant of either of the constituent parts of $-(\widehat{a/b})_j$, but the definition of $-(a/b)_j$, according to the model, includes μ_{sj} (equation 8.2.3) as a constituent part in its own right. Its significance in this (albeit incorrectly specified) equation is all the more interesting because it is not really a variable which could have been included as a determinant on the basis of ad-hoc reasoning.

1. Not shown in the table.

Paradoxically then, these two results do provide some justification for the decision to impose μ_{sj} and σ^2_{sj} in the regressions of the previous two sections.[1] At the same time, because of the greater approximation of mathematical form required by this alternative approach, a number of influences, which have been demonstrated by splitting up the speed of diffusion, now become blurred: notably, the role of $\sigma_{sj}{}^2$ and LI_j in the cumulative normal curve and μ_{sj} and π_j in the cumulative lognormal case.

A second test is whether Mansfield's model[2] would have been able to explain the data any better. In chapter 6, it was established that the logistic curve offered a slightly inferior description of the diffusion time paths. However, a good explanation of the variance of $\hat{b}_j$ from those equations may still be possible. In his own work, Mansfield found that π'_j (the pay-back from investing in the innovation, relative to the pay-back achieved on more typical investments in the industry) and S_j (the cost of the innovation relative to the total assets of the average sized firm in the industry) explained virtually all of the variations in $\hat{b}_j$ from 12 estimated logistic diffusion curves.

For the 22 innovations in this sample, data was not available for the typical investment pay-back and average assets of the adopting firms; therefore it is impossible to run exactly the same regression. However, equation 4 in table 8.6.1 is a close approximation, π_j is as defined above and $K_j/e^{\mu_{sj}}$ is used in place of S_j, where $K_j/e^{\mu_{sj}}$ is the cost of the innovation divided by the mean employment size of firms in the industry. As can be seen, neither of these variables approaches significance and a very low corrected R^2 is achieved. Even although the independent variables are not quite as Mansfield would wish, this collapse of explanatory power is quite startling. Equation 3 confirms that even using the 'best' explanatory variables from the earlier regressions, the fit is still mediocre,

1. The transformations of the dependent variables amount to imposing coefficients on $\widehat{\sigma^2_{sj}}$ and $(\widehat{\sigma/\beta})_j$ in equation (8.?.2) and $\hat{\mu}_{sj}$ in equation (8.2.3)
2. See chapter 2, section 2.

Table 8.6.1. Untransformed time series parameters as dependent variables.

Dependent variable.	Const.	$\log N_j$	$\log \Pi_j$	$\log \bar{C}_j$	σ^2_{sj}	$\log LI_j$	μ_{sj}	R^2	$\bar{R}^2$
Log $\hat{b}_j$ from the cumulative normal, n = 16									
1.	.929^{++} (1.01)	-.390 (4.53)	-.269^{+} (2.20)	8.134^{+} (1.98)	-.054^{++} (0.75)	.834^{++} (1.14)		.663	.495
log $-\widehat{(a/b)}_j$ from the cumulative lognormal, n = 13									
2.	.562^{++} (1.42)	.410 (2.32)	.309^{+} (1.94)				-.084 (2.24)	.706	.647
$\hat{b}_j$ from the logistic n = 22									
	Const.	N_j	Π_j	$\bar{C}_j$	σ^2_{sj}	K_j		R^2	$\bar{R}^2$
3.	.401 (6.45)	-.003 (4.50)	-.024^{++} (1.09)	+.024^{++} (1.31)	-.031^{+} (1.77)			.563	.460
4.	.295 (7.09)		-.011^{++} (1.23)			-.002^{++} (0.56)		.101	.006

Notes: figures in parentheses below estimated coefficients are t statistics.

+ denotes a coefficient not significantly different from zero at the 95% level.

++ denotes a coefficient not significantly different from zero at the 90% level.

with only N_j significant (although the negative coefficient on σ^2_{sj} is interesting.) The only explanation of these poor results must be that the logistic curve is not a good enough approximation to the diffusion data for these 22 innovations at least; consequently the estimates of b_j are rather meaningless and difficult to 'explain.'

Summary.

An overall evaluation of these results is rather difficult. On a purely statistical level, the only limited degrees of freedom available have led to spurious correlations between a number of explanatory variables which has prevented an unequivocal statement of their influence on the speed of diffusion. By the same token, of course, it would be wrong to attribute too much credence to those variables which have appeared as significant in the above analysis, given the smallness of the samples considered. This problem would have been ameliorated had it been possible to assume the same functional form for the diffusion curves of all the sample innovations; in that case, the cross-section analysis would have been based on 22 observations as opposed to the 15 and 13 used, respectively, for the cumulative normal and lognormal parameters. Given that this was not possible, the best solution, clearly, would be to increase the overall sample size in order that, say, 20 observations were available for each alternative curve. Having said this, it must be remembered that the only other systematic study along these lines to date is that of Mansfield, in which he used only 12 observations. Further, of these 12, 3. at least, may not have been appropriate, given the existence of autocorrelation in his time series regressions used to generate observations of the speed of diffusion (see section 2 of chapter 2.)

On the positive side, a fairly acceptable explanation has been possible for the two constituent parts of the cumulative lognormal speed of diffusion. The procedure of splitting the speed of diffusion into these constituent parts seems to have been justified, given the added precision possible, especially as compared to the regressions, reported in table 8.6.1, which ignore the nature of the relationship between the structural parameters and diffusion speed. Moreover, the results obtained are particularly impressive when compared with those generated by the logistic epidemic model.

On the negative side, the results reported in sections 4 and 5 suggest that there are large unexplained residuals in the constituent parts of diffusion speed - particularly for $(\widehat{6/\beta})_j$. This may be partly due to the problems of measurement of these constituents (especially $(\widehat{6/\beta})_j$), but also, it is probable that a number of important explanatory variables remain unidentified.

A discussion of the <u>economic</u> implications of the significance of N_j, π_j and LI_j (and, to a lesser extent, 6^2_{sj} and C_j) has been postponed until the following chapter.

<u>Appendix 1 to Chapter 8. Two sundry results.</u>

One seemingly anomalous conclusion of this chapter is that the determinants of the speed of diffusion are different for the two alternative growth curves. Thus, μ_{sj} determines the speed of the cumulative lognormal but not of the cumulative normal and vice versa for 6^2_{sj} and LI_j. It should be stressed, however, that this is due to the different measures used for the speed of diffusion and not because the structural parameters themselves have different explanations for the two curves.

This can be seen quite clearly when equation (8.3.19) is fitted for those innovations having cumulative lognormal diffusion. As for the cumulative normal set, LI_j is the only significant determinant[1]:

$$\log(\widehat{6/\beta})_j = \underset{(1.1)}{-.104} - \underset{(2.31)}{1.204} \log LI_j \qquad \bar{R}^2 = .192$$

Similarly, when (8.3.20) is fitted for those innovations having cumulative normal diffusion, only N_j, π_j, and $\bar{C}_j$ appear to be significant:[2]

1. cf. equation 10 of table 8.4.1.
2. cf. equation 1 of table 8.5.1.

$$\log\left\{-\left(\widehat{\frac{\log\alpha}{\Psi}}\right)_j\right\} = \underset{(2.75)}{9.154} \; \underset{(5.13)}{+.334\log N_j} \; \underset{(3.05)}{+.262\log \Pi_j} \; \underset{(2.15)}{-7.601\log \bar{C}_j} \; \underset{(0.66)}{-.058\,\sigma^2_{sj}}$$

$$\bar{R}^2 = .655.$$

Thus, each of $(\widehat{\Psi/\beta})_j$, $(-\widehat{\frac{\log\alpha}{\Psi}})_j$ and $(\widehat{\sigma/\beta})_j$ are determined by the same set of variables regardless of the form of the diffusion curve, the difference lies in the fact that $(-\widehat{\frac{\log\alpha}{\Psi}})_j$ does not influence the speed of diffusion for the cumulative normal curve and $(\widehat{\sigma/\beta})_j$ does not influence the cumulative lognormal speed of diffusion.[1]

Appendix 2 to Chapter 8: Cross-section analysis of $\hat{c}_j$ from the time series regressions.

In the main text of this chapter, no explanation has been offered for inter-industry differences in the estimated values of c_j given in Chapter 6. The reason for this is that using the 95% significance test, the null hypothesis of $c_j = 0$ can only be rejected for 5 of the 22 innovations. Crudely, then, there is very little to explain; 5 observations on c_j will obviously not support any rigorous cross-section analysis.

Out of interest, however, the tabulations presented in table (8.A2.1) were compiled, bearing in mind the hypotheses of section 6(V) of chapter 5.

Table 8.A2.1. Classification of innovations by $\hat{c}_j$.

	K(£000)	Π(yrs)	σ^2_s	N
Average value for innovations for which $\hat{c}$ is positive and significant at the 75% level.(7)	57.1	3.7	1.78	212
Average value for innovations for which $\hat{c}$ is insignificantly different from zero at the 75% level (8)	750.0	5.0	2.03	150
Average value for innovations for which $\hat{c}$ is negative and significant at the 75% level.(7)	885.0	3.7	2.4	74.

Source: table 6.6.1.

1. Strictly this is not true, $(\widehat{\sigma/\beta})_j$ influences $\hat{b}_j$ which in turn will affect $\widehat{SD}_{LN}$ as defined in (8.1.7.) In this chapter it has been convenient, however, to concentrate attention on $(\widehat{-a/b})_j$ as the main part of $\widehat{SD}_{LNj}$.

It should be stressed that no real significance may be attached to this table as it is based on a significance level <u>far</u> below any usually accepted level.

It is mildly interesting to note, however, that the average cost of innovations with a positive $\hat{c}$ is far less than those for which $\hat{c}=0$ (at the 75% level) which is, in turn, less than the average cost of innovations with negative $\hat{c}$. Had these results been based on 95% significance tests, it might have been worthwhile to use a χ^2 to test the hypothesis that, on average, more expensive innovations have a greater probability of being installed in periods of low capacity usage, whilst cheaper innovations have a greater probability of being installed in periods of high capacity usage.

There also appear to be slight tendencies for the positive estimates of c_j to occur in more competitive industries (lower σ^2_{sj} and higher N_j), but again, this is a strictly non-rigorous result.

Chapter 9 : Implications of the cross industry / innovation empirics, with special reference to the role of industrial structure.

This chapter is concerned solely with the implications of the results reported in the previous chapter: the implications of the empirical findings of the thesis as a whole are explored in chapter 10. Section 1 discusses the significant determinants of the cumulative normal speed of diffusion, as identified in chapter 8, and section 2 repeats the procedure for the cumulative lognormal speed of diffusion. The role for industrial structure which emerges from these two sections is somewhat complex, therefore, it is discussed separately in section 3. In section 4 the non-significant explanatory variables are considered briefly.

1. The cumulative normal speed of diffusion.

The cumulative normal speed of diffusion for the innovation in the j th industry has been defined[1] inversely by:

$$SD_{Nj} = \frac{\bar{z}_{2j} - \bar{z}_{1j}}{b_j} \qquad (9.1.1.)$$

where $\bar{z}_{2j}$ and $\bar{z}_{1j}$ are the normits of 2 yardstick levels of diffusion. As a specific example, consider the time lapse between diffusion reaching 5% and 95%, $\bar{z}_{2j} = 1.645$ and $\bar{z}_{1j} = -1.645$, in which case,

$$SD_{Nj} = \frac{3.29}{b_j} \qquad (9.1.2.)$$

where $b_j = (\Psi/\beta)_j \left((\sigma/\beta)_j^2 + \sigma_{sj}^2 \right)^{-1/2}$ (9.1.3.)[2]

From the previous chapter, excluding that set of innovations for which the cumulative normal is an obvious mis-specification,[3] and including only those

1. Section 1, chapter 8, particularly equation 8.1.10.
2. Equation 8.3.5.
3. That is excluding SP, ASB, SF, ADH, EH, CC, NCTURN: thus the results to be quoted are from part (a) of table 8.4.1. As the estimated coefficients are very similar when VD, SPC and SL are also excluded, the implications will be almost identical and will not be discussed here.

explanatory variables found to be significant, $(\widehat{\Psi/\beta})_j$ and $(\widehat{\sigma/\beta})_j$ may be 'explained' as follows.

$$(\widehat{\Psi/\beta})_j = e^{.799} N_j^{-.351} \Pi_j^{-.374} \qquad (9.1.4.)$$

$$(\widehat{\sigma/\beta})_j = e^{-.203} LI_j^{-1.331} \qquad (9.1.5.)$$

Substituting equations (9.1.3) - (9.1.5) into (9.1.2.),

$$(\widehat{SD})_{Nj} = 3.29e^{-.799} N_j^{.351} \Pi_j^{.374} \left(e^{-.406} LI_j^{-2.662} + \sigma^2_{sj}\right)^{\frac{1}{2}} \qquad (9.1.6)$$

which identifies 4 'variables' as being responsible for cross-industry differences in the speed of diffusion: the typical pay-back gained on adoption $(\Pi)_j$, the labour intensity of the industry (LI_j), the number of firms (N_j) and the inequality in their sizes (as measured by the variance of log. firm size, σ^2_{sj}).

The elasticity of SD_N with respect to each variable may be calculated as in table 9.1.1.

Table 9.1.1. Estimated elasticities. (cumulative normal.)

Explanatory variable (X_{ij})	$\frac{\partial SD_{Nj}}{\partial X_{ij}} \frac{X_{ij}}{SD_{Nj}}$	Value at the mean
N_j	$+.351$	.351
Π_j	$+.374$	.374
LI_j	$-1.331/(1+\beta_j^2\sigma^2_{sj}/\sigma^2_j)$	-.841
σ^2_{sj}	$+.5/(1+\sigma^2_j/\beta^2_j\sigma^2_{sj})$	.184

Evaluated at the mean values of $(\widehat{\sigma/\beta})^2_j$ and $\hat{\sigma}^2_{sj}$ (for the 15 innovations considered), SD_{Nj} is most sensitive to changes in LI_j and least sensitive to changes in σ^2_{sj}. As σ^2_{sj} rises and $(\sigma/\beta)^2_j$ falls, the elasticities with respect to these two variables converge, but, as can be seen, those for N_j and Π_j are constant.

Whilst the possibility of biassed estimates cannot be ruled out (particularly, perhaps, for LI_j in (9.1.5)), taken at face value these results provide considerable food for thought.

The elasticity with respect to LI_j is surprisingly large and even its sign is not totally expected. In section 3 (iiib) of the previous chapter it is argued that σ_j^2 will be influenced by LI_j in two (opposite) ways. First, where labour intensity is high, according to the vintage model argument, there is likely to be less heterogeneity in the operating conditions of the firms in the industry.[1] On the other hand, the more labour intensive is the industry, the less science-based is it likely to be and, so it was argued, the more heterogeneity are firms likely to show in their ability to assess new innovations. Apparently the former hypothesis seems better founded than the latter. Indeed, it is possible that the latter is the opposite of the truth: it may be that in craft-based industries, firms' abilities to assess are consistently poor[2] and that it is in the more science-based industries where large differences (and more heterogeneity) exist. However, no matter which of these hypotheses is most intuitively appealing, the sheer size of the elasticity is a little difficult to accept at face value. Clearly, it would be particularly interesting to see how this variable performed for a different set of innovations.

The elasticity with respect to Π_j is as expected and appears to consolidate Mansfield's and Griliches' findings.[3] A 1% increase in the typical pay-back gained from adoption leads to a .374% increase in the time lapse between 5% and 95% diffusion. That is a 1% decrease in typical profitability leads to a .374% decrease in the speed of diffusion. In terms of the model, therefore, the extent of and returns from search seem to be greater for more profitable innovations: non-adopters are under more pressure, are more likely to see the innovation as a solution to that pressure and find it easier to obtain information.[4]

1. See section 6(iii) of chapter 5 for an enlargement of this point.
2. Strictly speaking, this finding does not rule out the possibility that such industries are consistently better in assessment.
3. See section 2(b) Chapter 2.
4. See section 6(iv), Chapter 5.

The elasticities with respect to N_j and σ^2_{sj} are considered at some length in section 3. At this point it is merely noted that there are opposing implications regarding the influence of industrial structure: a reduction in both N_j and σ^2_{sj} appears to increase the speed of diffusion, yet whilst smaller N_j will be associated with increasing concentration, smaller σ^2_{sj} implies decreasing concentration.

All of the elasticities may be seen in a better perspective by returning to the precise measure of SD_{Nj} used here. For the 15 innovations considered, $\hat{b}_j = .15$, which implies that it will take 21.93 years for diffusion to increase from 5% to 95%. To shorten this time span by one year, any one of the explanatory variables would need to be changed as shown in table 9.1.2.

Table 9.1.2: the changes needed to speed up diffusion by one year.

	% change required.	Absolute change required.(at the mean.)
N_j	-13.02%	-10 firms.
σ^2_{sj}	-24.84%	-.865
π_j	-12.22%	-.551 years
LI_j	5.84%	+ 3% of value added to labour.

2. The cumulative lognormal speed of diffusion.

The time lapse between 5% and 95% diffusion with a cumulative lognormal growth curve is given by:[1]

$$SD_{LNj} = e^{-(a/b)_j}\left(e^{(1.645/b_j)} - e^{-(1.645/b_j)}\right) \qquad (9.2.1.)$$

However, this will not be used as the measure of diffusion speed in this section for reasons which become obvious when typical values of $(\widehat{-a/b})_j$ and $\hat{b}_j$ are inserted into the expression. For the 13 innovations in the 'cumulative lognormal acceptable set'[2], $(\widehat{-a/b})_j = 3.23$ and $(\hat{b})_j = .862$; with these values $SD_{LNj} = 180$. That is, 180 years must pass before diffusion reaches 95%! To all intents, then, on average, an innovation with cumulative lognormal

1. From equation (8.1.7).
2. See table 6.5.1.

diffusion will never diffuse completely. The reasons for this will be discussed in the following chapter, but on a purely mathematical level, it can be seen that this surprising result is due to 'innovation time'[1] slowing down drastically at a very early stage in diffusion.

In this case, therefore, a more meaningful measure of diffusion speed[2] might be the time lapse between 5% and 60% diffusion, whence

$$SD_{LNj} = e^{-(a/b)_j}\left(e^{(.253/b_j)} - e^{-(1.645/b_j)}\right) \qquad (9.2.2.)$$

where

$$-\frac{a}{b}_j = -\frac{(\log\alpha + \beta\mu_s)_j}{\Psi_j} = -\left(\frac{\log\alpha}{\Psi}\right)_j - (\beta/\Psi)_j\mu_{sj} \qquad (9.2.3.)$$

and

$$b_j = (\Psi/\beta)_j\left(\sigma^2_{sj} + \sigma^2_j/\beta^2_j\right)^{-\frac{1}{2}} \qquad (9.2.4.)$$

From the previous chapter, considering only the cumulative lognormal set of 13 innovations, and including only those explanatory variables found to be significant at the 95% level,[3]

$$-\widehat{\left(\frac{\log\alpha}{\Psi}\right)}_j = e^{.212}\, N_j^{.311}\, \Pi_j^{.273} \qquad (9.2.5.)$$

$$\widehat{(\Psi/\beta)}_j = e^{2.299}\, N_j^{-.406}\, \Pi_j^{-.234} \qquad (9.2.6.)$$

$$\widehat{(\sigma/\beta)}_j = e^{-.104}\, LI_j^{-1.204} \qquad (9.2.7.)$$

Substituting equations (9.2.3.) - (9.2.7.) into (9.2.2.) yields a rather complex expression, from which the elasticities of SD_{LNj} with respect to N_j, Π_j, LI_j, μ_{sj}, and σ^2_{sj} may be derived, as in table 9.2.1.

Table 9.2.1. Estimated elasticities (cumulative lognormal.)

Explanatory variable (X_{ij})	$\frac{\partial SD_{LNj}}{\partial X_{ij}} \cdot \frac{X_{ij}}{SD_{LNj}}$	Value at the mean.
N_j	$.311\left(-\frac{\log\alpha}{\Psi}\right)_j - .406\left[(\beta/\Psi)_j\mu_{sj} + (K/b)_j\right]$	.952
Π_j	$.273\left(-\frac{\log\alpha}{\Psi}\right)_j - .234\left[(\beta/\Psi)_j\mu_{sj} + (K/b)_j\right]$	1.239
LI_j	$-1.204(1 + \beta^2_j\sigma^2_{sj}/\sigma^2_j)^{-1}(K/b)_j$	-.430
σ^2_{sj}	$.5\ (1+\sigma^2_j/\beta^2_j\sigma^2_{sj})^{-1}\ (K/b)_j$	.067
$e^{\mu_{sj}}$	$-(\beta/\Psi)_j$	-.697

Footnotes from previous page:-

1. See chapter 5, section 5.
2. Note that in this chapter, the minor influence of b_j is also considered. It has been expositionally convenient in the previous chapter to refer to $-(^a/_b)_j$ as <u>the</u> influence on SD_{LNj}.
3. As shown in table 8.5.1. and Appendix 1 to the previous chapter.

where $K_j = .154 \dfrac{e^{\frac{1.898}{bj}} + 1}{.61\, e^{\frac{1.898}{b_j}} - 1}$ $= .53$ for the mean observed value of b_j.

It is quite noticeable that the elasticities with respect to N_j and π_j are both much larger, and those for LI_j and σ^2_{sj} both much smaller, than for the cumulative normal curve. However, in this case, the <u>sizes</u> of the elasticities are partly determined by the diffusion levels used in the computation of SD_{LNj}. For instance, had SD_{LNj} represented the lapse between 5% and 95% diffusion, the elasticities with respect to N_j, π_j, LI_j and σ^2_{sj} would all have been higher, particularly the latter two. On the other hand, had 5% and 50% been chosen, all four would have had lower elasticities and, indeed, for LI_j and σ^2_{sj}, they would have been equal to zero.

For this reason no detailed analysis of the magnitudes of these elasticities will be pursued, although it is suspected that the ordering of the elasticities of all but that for $e^{\mu_{sj}}$ is insensitive to the exact choice of diffusion levels.

The <u>signs</u> of the elasticities are all as for the cumulative normal which provides further confirmation of the hypotheses discussed in the last section. In this case, $e^{\mu_{sj}}$ appears as an extra determinant, by virtue of its role as a structural parameter of the model. The implications of this will be considered in the next section.

When mean values for $\widehat{(-a/b)}_j$ and $\hat{b}_j$ are inserted into expression (9.2.2.), it can be seen that the average time span required for cumulative lognormal diffusion to increase from 5% to 60% is 30.6 years. (Quite clearly, slower than along a typical cumulative normal growth curve.) To shorten this time span by one year would require a change in any one of the explanatory variables as shown in table 9.2.2.

Table 9.2.2. The changes required to speed up diffusion by one year.

	%age change required.	Absolute change required. (at the mean.)
N_j	-3.44%	- 5.06 firms
σ^2_{sj}	-48.80%	- 1.01
π_j	- 2.64%	-.078 years
LI_j	+ 7.60%	+3.98 % of value added to labour.
$e^{\mu_{sj}}$	+ 4.69%	+138.79 employees

3. The role of industry structure.

In appendix 6, it is argued that the best way to measure industrial structure, in the absence of much disaggregated data of relevance, is by the three parameters of the industry firm size distribution. N_j and σ^2_{sj}, the number of firms and the variance of the logarithms of their sizes, provide a representation of industrial concentration which is more flexible than the more conventional measures such as the concentration ratio or the Herfindahl index. Similarly, $e^{\mu_{sj}}$, the geometric mean[1] of firm size, is probably the best available measure of entry barriers.

As can be seen from tables 9.1.1. and 9.2.1., two of these parameters, N_j and σ^2_{sj}, have been found to be inverse determinants of the speed of diffusion both for the cumulative normal and lognormal curves. $e^{\mu_{sj}}$ is a positive determinant, but only for the cumulative lognormal. Whilst the estimated elasticities provide the first real evidence that industrial

1. Assuming a lognormal size distribution.

structure does have an influence on diffusion, there is no simple relationship such as 'greater competition leads to faster diffusion.'

Following the analysis of Appendix 6, reductions in the number of firms and increases in the mean firm size are both associated with <u>less</u> competitive industries - both changes would increase diffusion speed on the above evidence. On the other hand, a reduction in the inequality of firm sizes (σ^2_{sj}) would also increase diffusion speed, yet lower σ^2_{sj} are associated with <u>more</u> competitive industries.

Consequently, in this context, the conventional measures of concentration would be quite inappropriate for the analysis. Not only is there no monotonic relationship between diffusion speed and concentration measured in these ways, but also, at any given level for, say, the Herfindahl index, an increase in the value of that index may be associated with an increase or decrease in diffusion speed. Consider the following three cases: industries A and B both have the same number of firms ($N_a = N_b$), but the inequality in sizes of firms in A is greater than in B, ($\sigma^2_{sa} > \sigma^2_{sb}$). Therefore, ceterus paribus, diffusion speed will be faster in industry B. Almost certainly, the H index will have a greater value for A, indicating that it is more concentrated than B (which it is, of course.) However, in a third industry, C, if the number of firms is smaller than in A or B, ($N_b > N_c$), and the inequality in the sizes of firms is equal to that in B, ($\sigma^2_{sc} = \sigma^2_{sb}$) , then, ceterus paribus, diffusion speed will be faster in C than in B. Again, almost certainly, the Herfindahl index will have a greater value for C than for B, indicating that B is less concentrated than C [1]. Thus:

$H_a > H_b$ and diffusion is faster in B.

$H_c > H_b$ and diffusion is faster in C.

1. If firm size is lognormally distributed then $H = e^{\sigma^2_s} / N$ (see appendix 6), in which case, one can be certain that $H_a > H_b$ and $H_c > H_b$. Even in the absence of this assumption, however, it is likely that this conclusion will still hold.

Before considering the implications of this result, it is helpful to identify the economic reasoning behind these elasticities. It has been argued[1] that the level of competition may be a determinant of Ψ_j and α_j and, to a lesser extent, β_j and σ_j^2. In the regression analysis of the last chapter, however, of the three dimensions of competition, only N_j was found to be significant in the explanations of $(\widehat{\Psi/\beta})_j$ and $(\widehat{-\log\alpha/\Psi})_j$; the inclusion of $e^{\mu_{sj}}$ in these equations had no effect whatsoever and, although σ_{sj}^2 sometimes approached significance, its coefficient could not be established as different from zero at the 95% level.[2] It is possible that this lack of significance for μ_{sj} and σ_{sj}^2 is due to a cancelling effect (for example, that σ_{sj}^2 has a roughly equal influence on each of Ψ_j, β_j and $(-\log\alpha_j)$ and thus, when these parameters are expressed in ratio form, these influences cancel each other out.) But this would require a considerable co-incidence.

On the face of it, then, the only evidence is that greater concentration (through lower N_j) leads to higher Ψ_j and perhaps lower $(-\log\alpha_j)$[3], and thus to faster diffusion. However, there is an alternative hypothesis to account for the significance of N_j which seems more attractive, given the non-significance of the other dimension of concentration, σ_{sj}^2. It was argued above[4] that information diffuses, proportionately, at a slower rate, the more firms there are in an industry. In terms of the model, this would be reflected in lower Ψ_j.[5] If this alternative hypothesis is accepted, then it might be fair to conclude that there is no apparent causal link between concentration and any of the structural parameters. Nevertheless, an increase in concentration <u>so long as it is effected by a reduction in N_j</u>, should increase the speed of diffusion. This is a

1. See appendix 6, section 1 for a summary of these arguments.
2. Interestingly, the sign of σ_{sj}^2's coefficient was different for the two curves. Had these coefficients been significant, this would have implied that greater size inequality leads to higher Ψ_j and lower α_j for the cumulative normal but the reverse for the cumulative lognormal.

For footnotes 3, 4 and 5, see top of following page...

3. As N_j is not a significant determinant of $(\widehat{\sigma/\beta})_j$, it is unlikely that its influence on $(\widehat{\psi/\beta})_j$ derives from β_j.
4. Chapter 4, section 2(d).
5. And also higher α_j if the <u>initial</u> quality of information is poorer where N_j is large.

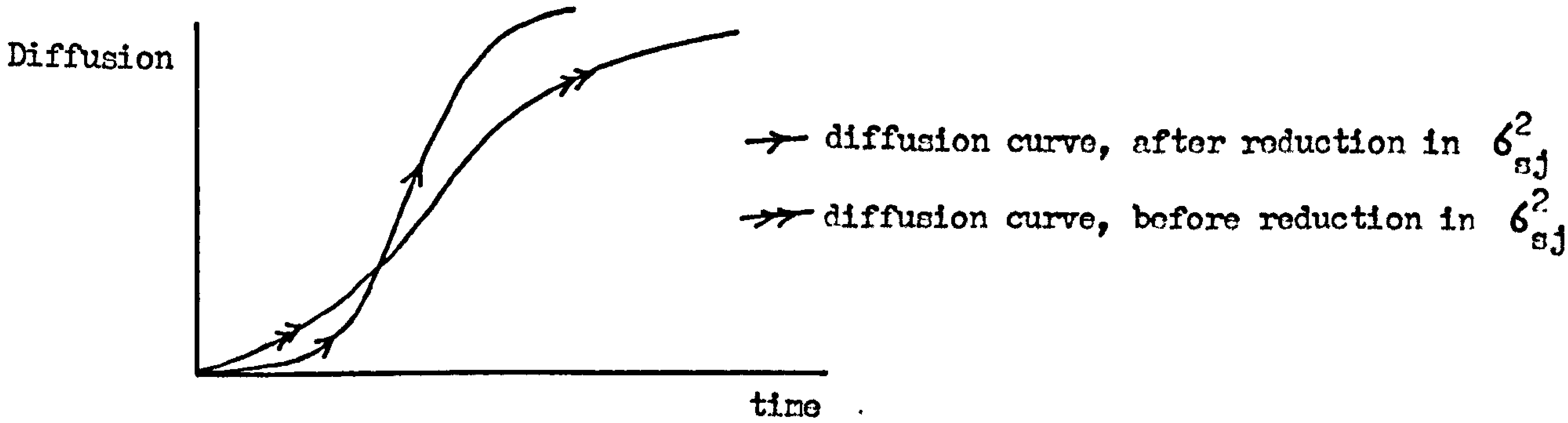
particularly interesting finding, as a new dimension is added to the long-running controversy on the effects of concentration on performance. Yet this is only the logical extension of the age old presumption that, where N_j is small, collusion is more likely as firms find it easier to get together and exchange information. In this particular context, however, information exchange is desirable, unlike in other areas such as pricing and output decisions.

Whilst σ^2_{sj} has been found to be an insignificant determinant of the structural parameters ψ_j, β_j, σ^2_j and α_j, it still sustains an influence on diffusion speed by virtue of its role as a structural parameter in itself. (Assuming that firm size does influence the speed of adoption of individual firms[1] and that firm size is lognormally distributed, then indisputably, the variance of log size <u>must</u> influence the speed of aggregate diffusion.) There are certain similarities here with the above discussion of N_j: higher σ^2_{sj} does not slow down diffusion <u>because</u> concentration, in itself, has detrimental effects. Rather, high σ^2_{sj} tend to occur in concentrated industries and a decrease in concentration <u>so long as it is effected by a reduction in σ^2_{sj}</u> should increase the speed of diffusion. In this model, high σ^2_{sj} reduces the speed of diffusion because it increases the heterogeneity of the population: not only is the profitability of the innovation more variable across firms, but also firms' ability to assess its value and their attitude to new innovations in general may be more variable.[2] Thus, any reduction in

1. And the various results of chapter 7 confirm this to be the case.
2. Some of this variability <u>may</u> be self-cancelling e.g. the innovation may be more profitable for large firms but, on the other hand, large firms may be more conservative about new techniques. There is no way of assessing this possibility. Nevertheless, as already noted the <u>net</u> effect is established in chapter 7 as increased firm size leading to a higher probability of adoption at each point in time (i.e. $\beta > 0$). Further, it can be shown that so long as $\beta > 0$ large firms will adopt earlier on average. (Appendix 3 to chapter 5.)

σ^2_{sj} must decrease this heterogeneity and, in consequence, reduce the importance of the two tails of the S shaped diffusion curve.

Figure 9.3.1. Effects of a reduction in σ^2_{sj} *



* assuming cumulative normal diffusion.

Similarly, the influence of $e^{\mu_{sj}}$ (in the cumulative lognormal case only) derives from the fact that μ_{sj} is, itself, one of the structural parameters of the model.[1] Because of the connection between firm size and both R^*_{it} and $(ER)_{Nit}$, the mean initial probability of adoption is higher, in any industry, the higher is mean size (and the greater are the benefits from large scale.) In the cumulative normal curve this really only influences diffusion in the very early years, but in the cumulative lognormal case this influence is more sustained. Thus an industry of large firms will adopt more rapidly because of the scale effect, but not because entry barriers are high per.se. (Assuming that industries with high mean firm size have high entry barriers as suggested in Appendix 6.)

Policy implications.

As has been indicated, the sizes of these three influences vary in the cumulative lognormal case, depending on exactly how diffusion speed is measured (namely, which yardstick levels of diffusion are used.) Consequently, for numerical simplicity, this discussion of the policy implications of these results will concentrate on the elasticities derived in the cumulative normal case.[2]

1. See section 2, equation 9.2.3.
2. For the cumulative lognormal the analysis is similar in essence, however.

On the face of it, the policy implications are straightforward: firm numbers should be reduced and the inequality in sizes of the remaining firms should be ironed out as far as possible. Two reservations need to be made, however. First, it would be wrong to extrapolate the results to values of N and σ_s^2 below those observed in the sample - the elasticities were estimated with only one exception (ATL), on data having a range, for σ_s^2, from .6 to 3.8 and, for N, from 20 to 584. Thus, there is not sufficient evidence to suggest that these elasticities would also apply to very high values or, more importantly, to very low values of σ_s^2 and N. The apparently logical conclusion - that a duopoly of two equally sized firms would be optimal in this respect cannot be made with any confidence.[1] Second, it seems very improbable that government policies to reduce N could be effected without, at the same time, increasing σ_s^2 and, possibly, it would also be difficult to reduce σ_s^2 without increasing N. Once these secondary effects of policy are admitted, the simplicity of the policy implications collapses.

<u>(i) The effects of a reduction in firm numbers.</u>

From Section 1, equation 9.1.6., the <u>total</u> effect on SD_N of a change in N may be defined as:

$$\frac{dSD_N}{dN} = .351 \frac{SD_N}{N} + .5 \frac{SD_N}{(\sigma^2/\beta^2 + \sigma_s^2)} \frac{\partial \sigma_s^2}{\partial N} \qquad (9.3.1.)$$

whence, $$\frac{dSD_N}{dN}\frac{N}{SD_N} = .351 + .5 \left(\frac{\partial \sigma_s^2}{\partial N} \cdot \frac{N}{\sigma_s^2}\right)\left((\sigma^2/\beta^2\sigma_s^2) + 1\right)^{-1} \qquad (9.3.2.)$$

From table 9.1.1, $$.5\left[(\sigma^2/\beta^2.\sigma_s^2) + 1\right]^{-1} = \frac{\partial SD_N}{\partial \sigma_s^2} \cdot \frac{\sigma_s^2}{SD_N} \qquad (9.3.3.)$$

1. Indeed, the only industry within the sample which approximates the pure oligopoly of theory - cars - records large residuals in the cross-industry empirics of the previous chapter, suggesting, perhaps, that the general effects of N and σ_s^2 do <u>not</u> apply at the extremely low observed values for these 2 variables.

Table 9.3.1. $\dfrac{\partial SD_{Nj}}{\partial \sigma^2_{sj}} \cdot \dfrac{\sigma^2_{sj}}{SD_{Nj}}$ for individual innovations.

Innovation.	
TK	.326
VM	.195
WSB	.130
NCTB	.141
ATL	.056
PCBC	.150
BOP	.277
TC	.203
GA	.060
CT	.126
F	.238
SF	.296
NCPP	.177
VD	.066
SPC	.314
Average	.184

and defining $\frac{\partial \sigma_s^2}{\partial N} \frac{N}{\sigma_s^2} = \eta_{\sigma^2 N}$, (9.3.4.)

then $\frac{dSD_N}{dN} \cdot \frac{N}{SD_N} = .351 + \eta_{\sigma^2 N} \frac{\partial SD_N}{\partial \sigma_s^2} \frac{\sigma_s^2}{SD_N}$ (9.3.5.)

Thus, whether a decrease in N will lead to an increase in the typical speed of diffusion ($\frac{dSD_N}{dN} \cdot \frac{N}{SD_N} > 0$)[1] will depend on the values of $\eta_{\sigma^2 N}$ and $\frac{\partial SD_N}{\partial \sigma_s^2} \cdot \frac{\sigma_s^2}{SD_N}$.

The mean value of the latter for the sample is given in table 9.1.1. as .184, but, as can be seen from table 9.3.1., for a number of innovations, the elasticity is well in excess of .184. It is quite impossible to generalise about the effects on σ_{sj}^2 of changing N_j – quite clearly, it will depend crucially on just how the change in N is effected. Usually, however, one might expect: $\eta_{\sigma^2 N} < 0$. That is, as N rises, σ_s^2 will fall.[2]

In table 9.3.2., 4 alternative values of $\eta_{\sigma^2 N}$ are assumed and against each one is listed the innovations for which a decrease in N would increase diffusion speed.

Table 9.3.2. The effects of a reduction in N.

Hypothetical value for $\eta_{\sigma^2 N}$				Diffusion speed increased for:
0	$\frac{\partial SD_N}{\partial \sigma_s^2} \frac{\sigma_s^2}{SD_N}$	: all values		i.e. all 15 innovations.
-1	$\frac{\partial SD_N}{\partial \sigma_s^2} \frac{\sigma_s^2}{SD_N}$	$<$	.351	i.e. " " "
-2	$\frac{\partial SD_N}{\partial \sigma_s^2} \frac{\sigma_s^2}{SD_N}$	$<$	.175	i.e. WSB, NCTB, ATL, PCBC, GA, VD.
-3	$\frac{\partial SD_N}{\partial \sigma_s^2} \frac{\sigma_s^2}{SD_N}$	$<$	.117	i.e. ATL, GA and VD.

1. Recalling that SD_N is an inverse measure of diffusion speed.

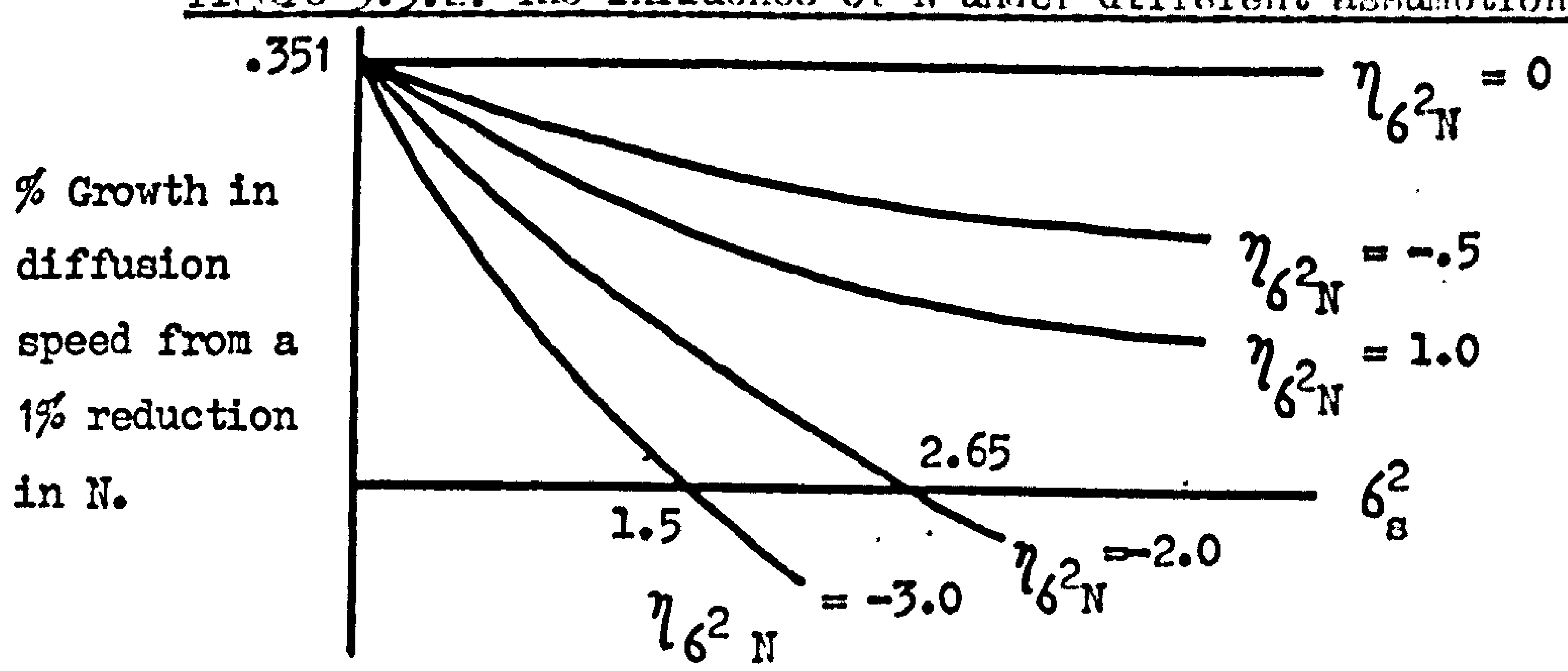
2. This is not to say that, in practice, industries with high N will have low σ_s^2 – this is certainly not true for the sample industries. The argument refers, rather, to the effects, in a given industry, of changing N holding other things, notably aggregate industry size, constant.

Thus, so long as $\eta_{\sigma^2_N}$ is inelastic, reductions in N will obviously increase diffusion speed. But for more elastic values, the policy will be counterproductive for a number of innovations. Indeed, for all values of $\eta_{\sigma^2_N} < -2$, a reduction in N will <u>decrease</u> diffusion speeds[1] for most innovations.

Similarly, the larger is $\frac{\partial SD_N}{\partial \sigma^2_s} \cdot \frac{\sigma^2_s}{SD_N}$, the more chance there is of a reduction in N being counterproductive, for given $\eta_{\sigma^2_N}$.

Now, from the definition of $\frac{\partial SD_N}{\partial \sigma^2_s} \frac{\sigma^2_s}{SD_N}$ in (9.3.3), this, in turn, implies that at larger values of σ^2_s, there is more chance of a reduction in N being counterproductive. In figure 9.3.2, the effects are shown of different values of $\eta_{\sigma^2_N}$ and σ^2_s, assuming a value for $(\sigma/\beta)^2$ of 4.9 (the mean for the sample.)

<u>figure 9.3.2: The influence of N under different assumptions.</u>



<u>(ii) The effects of a reduction in size inequalities.</u>

Much the same sort of analysis applies to policies aimed at reducing σ^2_s : the mean elasticity shown in table 9.1.1., or the individual ones in table 9.3.1, suggest that the diffusion speed[2] may be increased by acting to equalise firm sizes. In practice, however, this may only be possible by increasing the number of firms. Using similar algebra as

1. Or more correctly, a reduction in N would decrease the industry's propensity to diffuse innovations in the future. Obviously a change in N, now, will have little influence on the speed of diffusion of the sample innovations, given that they are all well into the diffusion process already.

2. Again, this is used as a shorthand for the industry's propensity to diffuse innovations in the future.

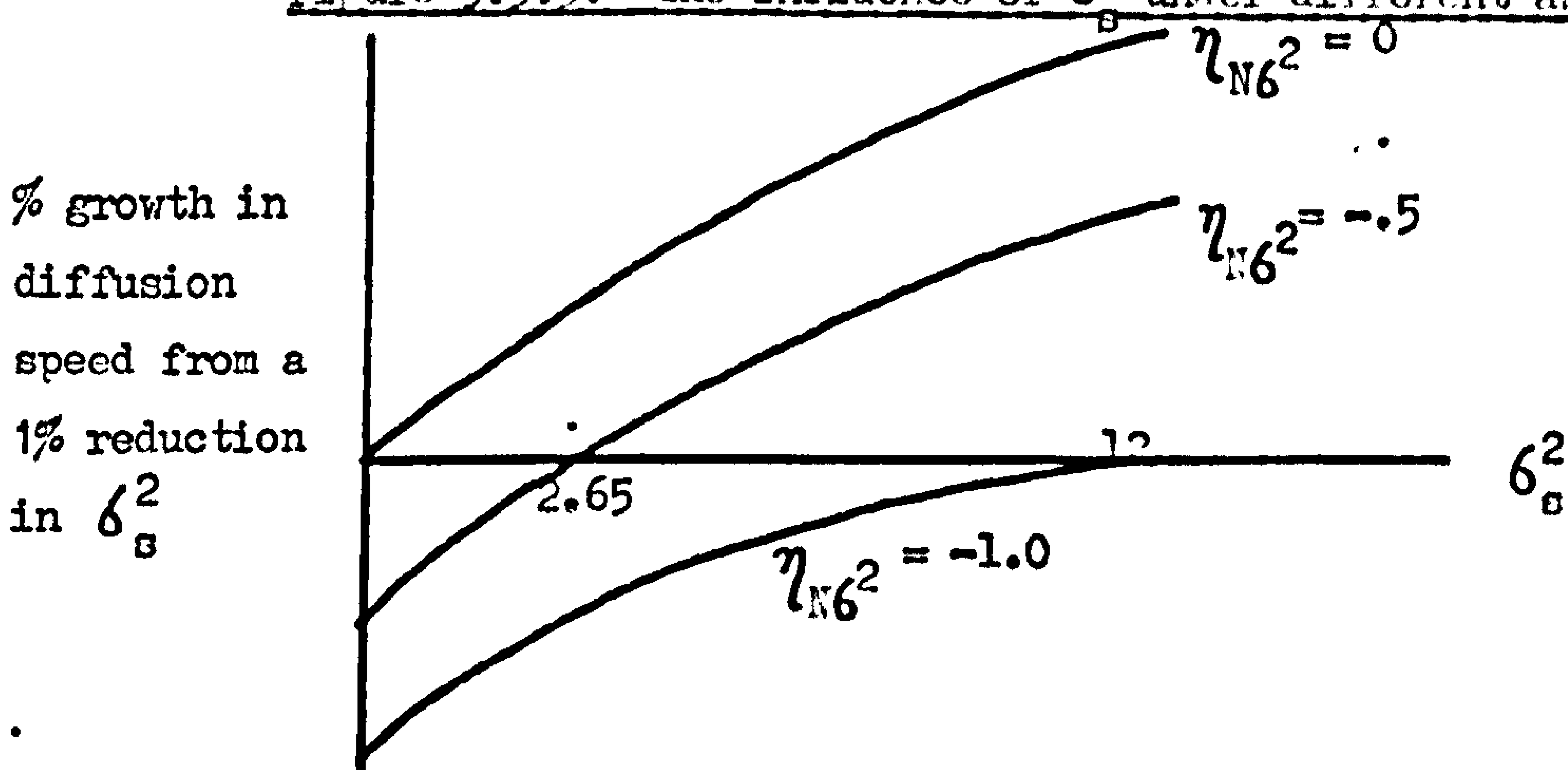
above, the <u>total</u> effect on SD_N of a change in σ_s^2 is given by:

$$\frac{dSD_N}{d\sigma_s^2}\cdot\frac{\sigma_s^2}{SD_N} = .351\,\eta_{N\sigma^2} + \frac{\partial SD_N}{\partial \sigma_s^2}\cdot\frac{\sigma_s^2}{SD_N} \qquad (9.3.6.)$$

$$\text{where } \eta_{N\sigma^2} = \frac{\partial N}{\partial \sigma_s^2}\cdot\frac{\sigma_s^2}{N} \text{ and } \frac{\partial SD_N}{\partial \sigma_s^2}\;\frac{\sigma_s^2}{SD_N} = .5\left[(\sigma^2/\beta^2\sigma_s^2)+1\right]^{-1} \qquad (9.3.7.)$$

Thus, the effects of a decrease in σ_s^2 will depend on the value of $\eta_{N\sigma^2}$ and of σ_s^2. Figure 9.3.3. shows the effects of reducing σ_s^2 for different values of σ_s^2 and $\eta_{N\sigma^2}$. (Again assuming $(\sigma/\beta)^2 = 4.9$.)

<u>Figure 9.3.3. The influence of σ_s^2 under different assumptions.</u>



Of course, if σ_s^2 can be reduced without increasing N, then, as can be seen from the diagram, given any reduction in σ_s^2, diffusion speed must increase. However, for the entire elastic range of values for $\eta_{N\sigma^2}$ and for all realistic values of σ_s^2, the effects of a decrease in σ_s^2 must be to <u>decrease</u> diffusion speed. In fact, using the observed values, $(\widehat{\sigma/\beta})^2$ and $(\widehat{\sigma_s^2})$, for each innovation and assuming only $\eta_{N\sigma^2} = -.5$, a reduction in σ_s^2 would actually slow down diffusion for seven innovations in the sample.

To summarise then, policies aimed at reducing N are less likely to increase diffusion speed (a) if in reducing N, σ_s^2 is increased and (b) if this is the case, the higher is σ_s^2 in the first place. Similarly, policies aimed at increasing diffusion speed by reducing σ_s^2 are less likely to succeed (a) if in reducing σ_s^2, N is increased and (b) if this is the case, the lower is σ_s^2 in the first place.

Reductions in N and σ_s^2 may be effected by a number of government policies. For instance, the former might be achieved by any of the following:

1. An active merger policy encouraging acquisitions between large and medium-large firms. This would certainly increase concentration and the market share of the leading firms would thus increase, without any reduction in the numbers of small firms. Under these circumstances $\eta_{\sigma^2 N}$ might be expected to be large, absolutely.

2. As above, but with the emphasis on take-overs of smaller firms by large or medium-large firms. Again this would increase the market share of the leading firms, but less so than in 1. In that the number of small firms is reduced, the variance of log size might not increase as much as in 1, and the size of $\eta_{\sigma^2 N}$ is less certain.

3. An active merger policy, encouraging mergers between small and medium-small firms. This would produce no increase in the market share of the leading firms. Overall, size inequalities might even decrease, as a larger proportion of industry output would be concentrated in firms of near average size.

In terms of the above analysis, alternative 1 would quite probably lead to a reduction to diffusion speed, alternative 2 might increase diffusion speed, so long as σ_s^2 was relatively low in the first place, and only alternative 3 would unequivocally lead to an increase in diffusion speed.

One other possible method of reducing N would be to:

4. Impose legal or institutional restrictions to discourage new entry and rely on natural wastage to reduce N.. In the short-run this might reduce the inequality between firm sizes as the number of small firms is reduced. In the long run, however, if the potential new entrants are 'more dynamic' than the existing population, this would impose less restraint on the internal growth of the larger firms at the expense of small firms, thus leading to an increase in concentration. With such a policy, therefore,

there would not necessarily be an increase in the industry's typical speed of diffusion.

Reductions in σ_s^2 might be effected by increasing the number of firms with near average size, at the expense of firms with extreme size at either end of the size distribution. This might be relatively difficult to achieve, using conventional policies. Two obvious possibilities, however, might be:

5. To encourage mergers between small firms (that is, as in 3, above), which, if successful, would not only reduce σ_{sj}^2 but also N_j. In such an eventuality, $\eta_{N\sigma^2} > 0$, which would definitely improve the chances for more rapid diffusion.

6. A policy of breaking up large firms. This might also reduce σ_{sj}^2, but only at the expense of increasing N_j. Such a policy would be most likely to have beneficial results if (a) concentration (and thus σ_{sj}^2) is high in the first place and (b) if the large firms were broken up into medium-sized, rather than small, firms.

In general, then, an industry's propensity to diffuse more rapidly may be enhanced by increasing the share of output of medium sized firms, at the expense of small firms; not only will this decrease size inequalities, but for a given aggregate industry output, N_j must also fall. As stated initially, this sort of policy implication may only be valid for the ranges of σ_{sj}^2 and N_j observed in the sample.

For the cumulative lognormal diffusion curve, the same sorts of conclusions apply. However, in this case, there is the added complication that increases in the mean size of firm also increase the speed of diffusion. Reductions in N_j must, fairly certainly, also increase $e^{\mu_{sj}}$; therefore, in this case, there is more chance that a merger policy of any kind would prove beneficial, but precise conclusions do depend on just how SD_{LNj} is measured. For instance, if it is defined as the time lapse between 5% and 95% diffusion, then the elasticities with respect to N_j and σ_{sj}^2 are much higher, relative to that for $e^{\mu_{sj}}$ than if the time lapse between 5% and 50% diffusion is used.

Thus, if one is concerned with reducing the duration of the later stages of diffusion, much the same conclusions may be drawn as for the cumulative normal. If the speed of the early stages is considered more important, then it matters very little whether policies increase the share of the larger firms, and thus σ_s^2, as the elasticity with respect to σ_s^2 will be relatively small.

One seemingly anomalous part of the analysis is that a decrease in the market share of large firms, under certain circumstances, might actually increase diffusion speed. This seems, at first sight, to be in conflict with earlier evidence that the larger is a firm, the earlier it will adopt. However, by definition, any decrease in the top firms' market share must increase the size of other firms in the industry (for a constant aggregate industry output.) To be sure, a decrease in the size of larger firms will slow down diffusion in the early stages, but it will approach 100% much earlier, as more of the curve is concentrated in the middle years of the period (see figure 9.3.1.) In the cumulative normal case then, overall diffusion speed is increased only at the expense of a slower take-off.[1] From the point of view of productivity, this may not be a bad thing, since substantial improvements in the new innovation are more forthcoming in the later years (given the form of the learning curve for Group B innovations.) The speed of the take-off should not be confused, moreover, with the <u>date</u> of innovation. The model does not predict that the first introduction of a new process will be earlier in an industry in which firms are large than in one typified by smaller firms. The process of innovation is taken as given, of course, in this

1. In the cumulative lognormal case, however, there is less chance that a reduction in the share of the leading firms will produce faster diffusion due to the probable negative effect on μ_s. But even here, reductions in the numbers of smaller firms would still help to reduce the characteristically long upper tail of the diffusion curve. It is this long upper tail, of course, which causes the typically lengthy time lapse before diffusion attains 95% (see section 2.)

context. In fact, from evidence to be presented in the next chapter,[1] it seems unlikely that the first introduction will be earlier for industries characterised by large size firms.

4. The non-significant determinants of diffusion speed.

Altogether, the model generated 9 measurable variables which might have been expected to influence diffusion speed. Of these, 4 variables ($\bar{C}_j$, VC_j, K_j and NS_j) proved to be consistently non-significant in all regressions at the 95% level and are now briefly considered.

$\bar{C}_j$ (the typical level of capacity usage in industry j throughout the diffusion period) was postulated as a possible determinant of firms' attitudes and the extent of their search (as represented by α_j and ψ_j). In the regressions, this variable had an estimated coefficient which was always consistent (but not significantly) with the hypothesis that high capacity usage encouraged more active search; moreover the size of the coefficient suggested that only small changes in $\bar{C}_j$ produced large changes in ψ_j. Indeed, for the cumulative normal, this variable was significant at the 90% level in some equations.[2] In this case the verdict must be 'not-proven' with some evidence to suggest that further investigation may be worthwhile.

VC_j (the variance of capacity usage in industry j) was used as a proxy for the underlying uncertainty in demand conditions: this might have been expected to have an effect on firms' investment yardsticks (as represented in α_j). This variable was not expected to provide substantial explanation and its total lack of significance is not too surprising. A query as to the appropriateness of this proxy must remain however.

K_j (the typical initial outlay required for adoption, also acting as a proxy for the technical complexity of the innovation.) The very poor performance of this variable in all regressions is a little surprising.

1. Chapter 10, section 2(a).
2. If there had been stronger, a priori, reasons for expecting a positive coefficient, then a one tailed test would have been justified. In that case, the estimated coefficient would have often been significant at the 95% level.

It was postulated as a determinant of all five structural parameters of the model and the arguments for its inclusion were often very strong. Because all three dependent variables $\widehat{(\Psi/\beta)}_j$, $\widehat{(\sigma/\beta)}_j$ and $\widehat{(-\log\alpha/\Psi)}_j$ are ratios, it is possible that K_j may still be a significant influence on each parameter individually, but that due to cancelling, it fails to show up as a determinant of any of the ratios. More probably, however, as specified, K_j is not an accurate measure of those influences (notably technical complexity.) It is difficult to conceive of any better quantifiable measures, but one alternative approach is possible, nevertheless. In table 9.4.1., the innovations are grouped into 4 technological categories and the mean values for $\widehat{(\Psi/\beta)}$, $\widehat{(\sigma/\beta)}$ and $\widehat{(-\log\alpha/\Psi)}$ are computed for each category. Definitions of these terms are provided in Appendix one; briefly, supplementary innovations are, typically, added to existing capital equipment and do not replace any existing equipment; vintage-model type innovations involve the replacement of an old technology embodied in existing equipment; new functions innovations entail the creation of a new sub-process in the existing overall process of the adopting firm and automating innovations are as usually defined. Technologically, one might expect each of these groups to be homogeneous in a number of respects (including, to a certain extent, cost and complexity.)

Table 9.4.1: The influence of technological characteristics on the structural parameters.

Innovation type.	Structural parameters.			
	$\widehat{(\sigma/\beta)}_j^1$	$\widehat{(\Psi/\beta)}_j^2$	$\widehat{(\Psi/\beta)}_j^3$	$\widehat{(-\log\alpha/\Psi)}_j^3$
Supplementary.	2.01(6)	1.72(3)	3.03(5)	5.56(5)
Vintage-model type.	1.96(8)	1.64(6)	1.56(4)	8.64(4)
New function.	2.39(2)	.45(2)	-	-
Automating.	1.69(6)	1.40(4)	1.67(4)	8.04(4)

N.B. Numbers of observations in each cell shown in brackets.

1. For all 22 innovations.

2. For those innovations for which the cumulative normal is not an obvious mis-specification.

3. For those innovations for which the cumulative lognormal is not an obvious mis-specification.

A simple-minded inspection of these mean values would suggest a number of noticeable differences, (for instance, that β may be higher, on average, for vintage-model type innovations.) However, the analysis of variance, which takes into account the intra- and inter-group variances around these means, yields F values which show that there are no significant inter-group differences in any of these 4 ratios.

NS_j : (the number of firms supplying the innovations.) This was used as a very crude proxy for the level of competition in the supplying industry. Its lack of significance cannot be attributed to any cancelling effect, as it was postulated as a determinant of Ψ only. However, more rigorous tests are needed before a definitive statement is possible on the role of structure in the supplying industry. As an alternative test, the analysis of variance was used to test the hypothesis that $(\widehat{\Psi/\beta})_j$ for innovations supplied by monopolists were significantly different from $(\widehat{\Psi/\beta})_j$ for all other innovations, but no significant difference emerged.

Summary.

A number of points of interest have emerged from the results of chapter 8. First, in common with the major finding of past research, the profitability of the innovation is clearly an important determinant of the speed of diffusion, whether along a cumulative normal or lognormal growth path. Second, and somewhat surprisingly, labour intensive industries will tend to have faster diffusion, ceterus paribus, than more capital intensive industries. Whilst this result may be easily rationalised by equating labour intensity with homogeneity of operating conditions in the potentially adopting population of firms, the estimated elasticity of this variable is surprising. More research on this particular result is surely desirable. Third, the cost of the innovation appears to have no influence on the speed of diffusion, as opposed to a weakly significant result to the contrary of

Mansfield's. Fourth, the industrial structure of the consuming industry is clearly of some importance, but its influence on diffusion speed is particularly complex.

Finally, an interesting implication, which has emerged in passing, is that diffusion along a cumulative lognormal growth path will, typically, slow down so quickly that 100% diffusion is never attained.

Chapter 10 : Implications, Summary and Conclusions.

This chapter falls into three distinct parts. First, the broad implications of all of the empirical results of the thesis are considered. Partly to add some perspective to these results, section 2 attempts an assessment of the successes and limitations of the theoretical model and the empirical relationships it suggests. Finally, section 3 considers one or two developments which might be made and which provide interesting implications for the study of economic growth.

1. Implications.

The more interesting empirical findings of this thesis may be grouped into three broad areas: the influence of firm size and industrial structure, the effects of demand conditions facing the consuming industries and the role of the technological characteristics of the innovations concerned.

(a) Firm size and industrial structure.

Much of past research into the determinants of technical progress in general has concentrated on the role of firm size and industrial structure. A partial concencus appears to have emerged.[1] Reverting to the invention-innovation-imitation classification outlined in the introduction, this concencus may be summarised briefly. The propensity to invent (as measured by patents and/or R & D. expenditures) appears to rise with firm size but usually at a slower rate proportionately; industrial concentration tends to be mildly positively correlated with inventive activity, but the causal direction is by no means certain. Turning to innovation, in what is generally agreed to be the most exhaustive study to date, Mansfield [2] finds that the number of major innovations made by firms rises as a cubic function of their size. Williamson,[3] using Mansfield's data, finds that the ratio of the top four firms' share of

1. See L. Weiss, "Quantitative studies of industrial organisation", pp.389-397, in "Frontiers of Quantitative Economics," ed.M.Intriligator, North Holland, Amsterdam, 1971. and Kennedy and Thirlwall, op.cit. pp.43 - 62.
2. Mansfield, (1968) op.cit. chapter 5.
3. O. Williamson, (1965) op.cit.

innovations to their share of capacity declines as concentration increases. The only findings in the area of diffusion have already been discussed in chapter 2. To recapitulate, Mansfield and others find that, on average, the length of time a firm waits before adopting is inversely related to its size. Further, Mansfield's estimates of the speed of diffusion for 12 innovations (but in only 4 industries) are not inconsistent with the hypothesis that diffusion is faster the lower is industrial concentration.

In this thesis, it has been found that industrial structure does have an influence on diffusion speed, but in such a way that it is impossible to say, for instance, that high concentration always leads necessarily to faster or slower diffusion. As the discussion of the previous chapter [1] is fairly self-contained there is little point in any repetition here. On the other hand, a brief summary of the main findings on firm size is provided, if only because a number of important conclusions have emerged at different stages in the empirics and an overview is helpful.

First, however, as a digression, the data collected on the sample innovations does provide the opportunity to test the relationship between firm size and the propensity to innovate.[2] For each of the 18 sample innovations for which data is available, table 10.1.1 shows the size of the innovating firm. As can be seen, in only 6 of the 18 is the innovator larger than the arithmetic mean sized firm in its industry. In this context, however, the median size is perhaps more meaningful: if, in each industry, <u>each firm has an equal propensity to innovate</u>, then one might expect 9 of the 18 to have been innovated by firms with size in excess of their industry median. In fact for 13 of the innovations this is the case. Apparently, the larger halves of these industries have innovated more than the smaller halves.

More interesting perhaps, is whether large firms do 'their share' of innovating, given their share of total employment. This would be true if <u>each firm has a propensity to innovate which is proportional to its share of industry employment</u>. There are a number of ways of testing this hypothesis;

1. Section 3 chapter 9.
2. In the sense of being the first adopter of the new process.

in the event, a relatively simple test is used.

In each industry, firms have been ranked in order of size and then categorised into four groups: 'large' firms, 'medium-large', medium-small' and 'small' firms. Each group accounts for one quarter of the total industry employment. For each innovation, the group which includes the innovating firm is shown in column 4 of the table.

On this basis, quite clearly, large firms have not accounted for their share of innovating: for only one innovation is the innovator a large firm, 3 are medium-large, 3 small-medium and 11 are small firms. For the record, a χ^2 test at the 5% level rejects the hypothesis that these differences are not significant.

To return to the main theme of this thesis, however, the model generates two alternative specifications of the firm-size adoption-performance relationship. The Quasi-Engel curve describes the relationship between probability of adoption and firm size at any point in time. Alternatively, the delay before adoption can also be shown to be a simple function of firm-size.[1]

This second relationship has not been tested directly for the sample, due to the problem of bias which arises because diffusion is not yet complete for most of the sample innovations.[2] However, it can be shown that this relationship adds no information which could not be deduced from the Quasi-Engel curve and the diffusion curve parameters.[3] In passing, it has been noted that Mansfield's use of the adoption-delay versus firm-size relationship seems to be inconsistent with his logistic diffusion curves.[4]

Due to the limited degrees of freedom with which the Quasi-Engel curves have been estimated, it would be wrong to be too definite in the conclusions drawn, but on the face of it, they seem to establish fairly conclusively that firm size is an important determinant of behaviour in this context.

1. Appendix 3 to chapter 5.
2. See the discussion of this point in section 3(b) of chapter 2.
3. See the note on ATL in Appendix 4.
4. See Appendix 3 to Chapter 5.

Table 10.1.1. The size of innovating firms.

Innovation	Size[1] of first adopter (employees)	Arithmetic mean for industry	Median size for industry	First adopter's size class
NCTURN	6500[2]	614	283	Med-large
SPC	6653[2]	1048	738	Large
SL	750[2]	168	40	Med-large
NCTURB	5000[2]	1686	992	Med-small
GA	110[2]	45	22	Med-large
NCPP	1200[2]	527	186	Med-small
BOP	8400	8628	5884	Med-small
WSB	700[2]	768	172	Small
VM	1100	1336	665	Small
F	429[2]	768	172	Small
SF	429[2]	768	172	Small
PCBC	323	768	172	Small
SP	212[2]	768	172	Small
CT	325	462	337	Small
TK	27[2]	98	34	Small
VD	1500	5714	3213	Small
ATL	10401	27536	25360	Small
CC	350	7609	4675	Small

1. Size in the most recent year for which data is available, not at the time of adoption.

2. First adopter in the sample of firms for which data is available. See Appendix 2 for details of these samples.

For definitions of large, medium-large etc., see text.

A visual inspection of the empirical curves, backed up by χ^2 tests, confirms that for all of the sample innovations, probability of adoption is, indeed, positively related to firm size. Moreover, in no case is the data inconsistent with the hypothesis that the relationship follows the cumulative lognormal curve suggested by the model.

As an interesting extension, it is possible to use the parameters of the estimated Engel curves to assess the importance of size as a determinant of $(R^*/ER_N)_{ijt}$ - the ratio which is fundamental to the construction of the theoretical model in Chapter 5. In all cases, size is shown to be a significant determinant and for 8 innovations it explains half or more of the variance between firms in this ratio.[1]

It should be stressed, however, that these results do not establish large firms as being more progressive. The survey of the technical characteristics of the sample innovations and industries shows that there are a number of pervasive technical reasons.[2] why most new processes should be more profitable for larger firms. Thus, it would be surprising if there was not a positive relationship between probability of adoption and firm size. As it is not generally possible to quantify these technical advantages very accurately, the results cannot establish whether, for instance, large firms are better information gatherers and/or are more favourably disposed to new innovations.

Finally, on firm size, Weiss[3] chides Mansfield for not attempting to use his results (which are, of course, consistent in broad terms with those presented in chapter 7,) to answer 'the more basic question of whether large firms were quicker to imitate (adopt) than groups of smaller firms with about the same number of investment decisions to make.' On reflection, there seems

1. See column 6 of table 7.2.1. R^*_{it} is the yardstick used by firm i to assess whether adoption of innovation j is acceptable, ER_{Nijt} is its expectation of the profitability to be gained from adoption, both at time t.
2. See section 6 of chapter 3.

3. op.cit., p.396.

to be little point in attempting such an analysis, however. For instance, if it <u>were</u> shown that, say, one large firm with 20,000 employees in the steel industry adopted earlier than did the average of 10 small firms, each with 2,000 employees, <u>but</u> at least one of these 10 smaller firms adopted earlier than the large firm, the policy implications would be unclear. Certainly, this would not provide enough evidence to support splitting the larger firms into 10 parts. The objective, presumably, is to enable <u>as much</u> of the industry output to be produced, by <u>as soon</u> as possible, using the new process. Yet in the above example the 20,000 employees in the large firm have access to the new process <u>before</u> even 10,000 of the employees of the smaller firms (abstracting from intra-firm diffusion, in both cases.)

As it happens it would be possible to use the parameters estimated in chapter 7 to answer Weiss's point.[1] But because of the preceding argument this possibility is not pursued.

<u>(b) The effects of the trade cycle and demand conditions.</u>

In that the adoption decision at the firm level may be viewed as only a special form of the investment decision, one might expect that the state of demand for the potential adopters' output would play a significant role in the diffusion process. Indeed, as was reported earlier, Gold et al[2] provide some (non-conclusive) evidence that fast growth in demand is a necessary, but not sufficient, condition for rapid diffusion. In the other areas of technical progress, Schmookler[3] finds that <u>inventive</u> activity is at least partially determined by demand conditions but Mansfield finds no support for the hypothesis that <u>innovative</u> activity varies over the trade cycle.[4] In the theoretical model of chapter 5, demand conditions were hypothesised as having two main influences.

1. As an example, the probability that a firm of size S will adopt before one of size 2S can be computed from knowledge of (δ/β). Then, if the probability is greater than 1/4, the implication is that, on average, one of two firms of size S will adopt more quickly than a firm of size 2S.

2. See section 2 of chapter 2.

3. J. Schmookler, "Invention and economic growth", Cambridge (Mass.) Harvard University Press, 1966.

4. E. Mansfield (1968) op.cit.

First, at the peak of their trade cycle, adoption should be more beneficial because the adopting firm may be better able to run the new (often continuous) process at near full capacity; on the other hand, the disruptive influences of adoption are likely to be more costly if demand is high. Second, it was postulated that firms' efforts to search for more information about new processes will be influenced or even caused by their external environment, although it was not specified whether buoyant demand for their product would discourage or encourage potential adopters' search efforts. The empirical extensions of these arguments are twofold: demand conditions might influence not only the shape of the time series diffusion growth curve[1] but also its parameters. In fact, the empirical findings provide only sketchy support for either hypothesis. For a minority of innovations, a significant cyclical component, superimposed on the underlying S shape, is apparent. In some cases diffusion appears to have been speeded up in times of high capacity and retarded when capacity was low, but for others the reverse is true. As this effect is only significant for a third of the sample,[2] however, any systematic explanation of its size and sign is impossible. Because of the poor quality data used it would be wrong to read too much importance into this result, one way or the other; the inevitable conclusion must be that more research is needed in this area.

In some ways the second hypothesis - that the parameters of the growth curve may be affected by the level of economic activity - is the more fundamental from the policy viewpoint.[3] If it could be shown, for instance, that diffusion speed, as represented by the parameter b_j, is positively related to the level of economic activity, then, through demand management, the government would have at least some potential control over the pace of technical progress.

Unfortunately, such a result is not forthcoming using the usual 95% significance test. For the cumulative normal growth curves, $\bar{C}_j$ (the average level of capacity usage in industry j over the diffusion period) is found to be a positive determinant of $(\widehat{\Psi/\beta})^4_j$ - but this result is only significant at the 90% level.

1. See figures 5.5.5., for instance.
2. See table 6.6.1.
3. Cyclical fluctuations around the trend growth have no effect on the long-term speed of diffusion, after all.

(Footnote continued from previous page:)
4. This relationship, if significant, would identify $\bar{C}_j$ as a positive determinant of diffusion speed in this case.

For the cumulative lognormal growth curves, $\bar{C}_j$ is a positive determinant of $(\widehat{\Psi/\beta})_j$ and a negative determinant of $(-\widehat{\log\alpha/\psi})_j$ [1] but in neither case significantly so, even at the 90% level.

Once again, therefore, the question remains unresolved, these results are perhaps sufficiently pervasive not to completely rule out a connection between high capacity usage and more rapid diffusion, but their lack of statistical significance prevents any definitive statement. It is tempting to suggest that other measures may have been more appropriate, but, as is so often the case at this micro-level, severe constraints are imposed by the lack of adequate data.[2]

(c) The characteristics of innovations.

Somewhat ironically the strongest findings and conclusions to have emerged from the empirical side of this thesis have occurred in this area. It was intended that the examination of the technical aspects of new innovations should act merely as a backcloth to discussion of other factors, such as the roles of industry structure and demand conditions in the diffusion process. However, at least three sets of conclusions deserve some emphasis, in their own right.

First, and fairly unequivocally, a major part of the differentials between innovations in the speed at which they diffuse may be attributed to differences in their profitabilities. Not only is profitability robustly significant in the cross-section empirics of chapter 8 but also, the typical size of its co-efficient is fairly high for both cumulative normal and lognormal diffusion. This is not, of course, a new finding: both Mansfield and Griliches[3] have found a similar relationship.

1. Both relationships, if significant, would identify $\bar{C}_j$ as a positive influence on diffusion speed.

2. See Appendix 3.

3. See chapter 2 section 2b.

This new evidence really only serves to establish the generality of this result to a wider range of innovations and a different country, the U.K.

One might conclude that this suggests a certain measure of rationality on the part of the entrepreneurs, but it certainly does not establish profit maximising behaviour[1] - at least not in its more simple forms. Indeed, it is very difficult to reconcile the duration of the typical diffusion process[2] with any other than substantially modified theories of profit maximisation. Rather, it has been argued, here, that entrepreneurs will tend to be more active in their search efforts, the more profitable is the new process innovation. Not only do highly profitable innovations produce greater competitive pressures on non-adopters, but also they are more likely to be seen as the potential solution of various problems as they arise.[3]

Secondly, there does seem to be a marked difference in the shape of the diffusion growth curve between two types of innovations. As table 6.7.2. establishes, with only one or two exceptions, it is possible to predict that large, expensive and technically complex innovations (termed Group B) will have symmetric, cumulative normal diffusion whilst small, cheaper and relatively simple innovations (Group A) will tend to diffuse along a non-symmetric, cumulative lognormal curve. For innovations which do not fit easily into one or other of these classifications, not surprisingly, it is difficult to establish conclusively that their diffusion curve is normal or lognormal. This, incidentally, helps to reconcile two apparently contrary findings in past research. Mansfield, studying the diffusion of 12 major innovations in the U.S., finds that the symmetric logistic curve (not unlike the cumulative normal in its shape) provides an adequate description. Metcalfe, on the other hand, finds that the heavily skewed logarithmic reciprocal transformation (quite similar to the cumulative lognormal) provides

1. As implied by some commentators, see the quote on p.13 of chapter 2.
2. See the introduction to chapter 4, but also Mansfield(1968), op.cit., for evidence that it is not only in the U.K. that diffusion is typically a long drawn-out process.
3. See chapter 4, sections 4 and 3(c).

a fair fit to the observed diffusion of 3 innovations in the U.K.[1] Quite clearly, the technical descriptions of these innovations provided by both authors establish that, with one exception, Mansfield's 12 innovations are all Group B and Metcalfe's all Group A.

The rationale given for this difference between the innovations lies in the different sorts of learning curves which might be expected.[2] It is argued that with Group A innovations, the scope for learning, mainly by their manufacturers, diminishes fairly rapidly after quite substantial improvements in the early years. For Group B innovations, however, learning will be more sustained, with less tendency to tail-off, at least for most of the diffusion period. As a consequence, it seems probable that both competitive pressures on non-adopters and the returns from continued information search will continue for longer periods in the case of Group B innovations; whilst at a relatively early stage, further search concerning Group A innovations will produce little extra information and similarly, competitive pressures will stabilise.

One interesting implication of this hypothesis for Group A innovations is that, at a relatively early stage in diffusion, the propensity for non-adopters to adopt may decline markedly. If, after the initial stage in diffusion, non-adopters are able to withstand the competitive pressures[3] and are fully aware of the capabilities of the new innovation, there is, presumably, less chance of a later change of mind - especially if the innovation itself is subject to only minor further improvements in specification. Thus, it is not surprising to find that for the sample

1. Both studies are discussed in more depth in section 2 of chapter 2.

2. Section 7 of chapter 3 and also section 5 of chapter 5. Post-estimation scrutiny of the two sets of innovations and industries suggests no other obvious differences between them.

3. Because most Group A innovations are relatively small, the relative cost advantage for adopters will also be small and may well be swamped by other efficiency differences. This is much less likely for the typical Group B innovations, e.g. Tunnel Kilns, which are far more costly, but which also provide substantial reductions in operating costs.

Group A innovations, the parameters of the estimated diffusion curves suggest that diffusion will not approach 100% within the forseeable future.[1]

Incidentally, this distinction between innovations does answer the criticism that is sometimes made of the existing models of diffusion, namely that they do not differentiate analytically between major and minor innovations.[2]

Finally, and more generally, this apparent difference in the shape of the learning curve does also add another dimension to the conventional literature on learning by doing.[3] The examples usually quoted in this literature are equipment, produced off-site and in relatively large quantities, e.g. standard machine tools, airframes etc. Moreover, these products are fairly homogeneous. On the other hand, the typical process identified here as Group B tends to vary widely in its specification, depending on the needs of the consumers. Thus, initially, learning is far less likely to be transferable from one installation to the next. Only after a number of installations have been completed is the manufacturer able to consistently duplicate past installations and draw upon learning. Having said this, it is suspected that even for Group B innovations, learning will slow down rapidly after some stage, although this stage may only be attained when the imitation process is nearing completion.[4]

The third set of conclusions in this area arise from the technical descriptions provided in Appendix one and summarised in chapter 2. In general, it has been possible to describe the sample innovations by a number of characteristics which have been used, where possible, in the development of the model. In addition, however, these conclusions may be of more general interest.

The typical new process originates from outside of the industry in which it is to be used. This suggests only limited applicability for those theories based on the assumptions that the new process is invented and innovated by a firm intending to use it for its own purposes, and who may or may not permit its competitors to imitate the process under licence. Further, new processes

1. See section 2 of chapter 9.
2. See Professor Oshima's comment in Williams (1973), pp.282-3.
3. For instance, K. Arrow (1962) op.cit.
4. See chapter 3, section 7.

are invariably mainly labour saving often with fixed coefficients of production. In conflict with the implicit assumptions of the epidemic model[1] are the widely observed facts that profitability of adoption varies quite considerably across firms and time. Much of this variance across firms is due to the undoubted benefits accruing to large scale adoption. Indeed, economies of scale seem to be an essential feature of many of the sample innovations.

It is also interesting to note that many of the innovations produce significant changes in the quality of the product to be produced. This raises the definitional problem of whether 'process innovation' is a satisfactory nomenclature for the sample innovations studied.

Finally, for a majority of the innovations, the technological assumptions of the simple vintage model seem inappropriate. More than half of the innovations do not require the replacement of a technology embodied in existing capital equipment and, for those that do, the assumption of indivisible plant is violated. Admittedly this second assumption is not crucial to the vintage model but its relaxation does destroy the analytical simplicity of that model.

2. An overall evaluation of the model.

Gold, Pierce and Rosseger have argued that the large number of special, usually technical, factors make the construction of a general model of diffusion pointless and perhaps misleading.[2] This has always been a difficult argument to accept given the empirical success of the simple models used by Mansfield and Griliches. In this thesis, using a similar empirical methodology to Mansfield's, the data has surely provided further justification for refuting Gold et al's assertion.

In addition, the extra complication of the theoretical model has provided a more satisfactory theoretical link between the three areas of empirical analysis: curve fitting, explanation of inter-industry differences in the speed of diffusion and the relationship between individual firms' behaviour

1. See section 1, chapter 2.
2. See section 2 of chapter 2.

and their sizes.[1]

Moreover, whilst the model is based upon a number of fairly strong assumptions, where these have been testable they have not appeared to be unreasonable. In particular, three assumptions have been largely vindicated by the data. First, the hypothesis that firm size is lognormally distributed can not be rejected for the sample industries.[2] Second, the hypothesis that firm size is a major determinant of ER_{Nit} and R^*_{it} may be accepted for all of the sample industries from the analysis of the estimated parameters of the Quasi-Engel curves.[3] Strictly speaking, this assumption is not crucial for the time series implications, but without it the model would become something of an empty box. Third, the assumption of a lognormally distributed error term in the expression:

$$(R^*/ER_N)_{it} = \alpha\,\theta_t S_i^{\beta} \epsilon_i \qquad \text{(equation 5.2.7)}$$

is consistent with the data,[4] on the other hand the data used to test this hypothesis is sufficiently rough to rule out any rigorous tests.

Similarly the empirical success of the model is fairly encouraging. Most notably, the curve-fitting exercises of chapter 6 suggest that the high level of explanation achieved in the past by the logistic may be bettered, in this instance, by the use of the cumulative normal or lognormal curves. For only three of the twenty-two sample innovations is there evidence that both of these alternatives are mis-specifications of the diffusion curve. The estimated Quasi-Engel curves of chapter 7 confirm that there is also a fairly strong relationship, as predicted, between individual firms' probability of adoption and their size. Further, the cross-industry and cross-innovation empirical work of chapter 8 suggeststhat the model can explain a large portion of the observed differences in the speed of diffusion. The explanatory power of the model in this last context is roughly on a par with those results

1. On an entirely different level, a common theoretical thread has also been established with the models of diffusion of new consumer durables (see chapter 2 section 6.)
2. Strictly speaking, the lower tail of the observed size distribution of one industry (bricks) shows a significant deviation from lognormality. But in this case, this may be rationalised by a bias in data collection.
3. Chapter 7, pp. 5-6.
4. Chapter 7, pp. 7-8.

reported by Griliches but not as impressive as Mansfield's. However, Mansfield's estimating equations do not provide a good explanation of the rate of diffusion for the innovations in the present sample.

There are a number of ways in which the model might be developed to provide a more satisfying theoretical description. Two aspects in particular should perhaps be mentioned.

In chapter 5, firm size is the only firm-level characteristic <u>identified</u> as determining R_{Nit}, H_{it} and R^*_{it} [1]; the error term in equation (5.2.7.) thus represents all other unidentified determinants. Quite clearly there must be a wide range of factors, both technical and economic, which will influence firms' behaviour. For instance, this relationship could be extended as follows:

$$(R^*/ER_N)_{it} = \alpha S_{it}^{\beta} E_{it}^{\gamma} X_{it}^{\delta} \theta_t \epsilon'_{it} \qquad (10.2.1)$$

where E_{it} might be the educational attainment of the managing director of firm i and X_{it} some relevant technical characteristic such as, in the case of numerically controlled machine tools, the proportion of firm i's output produced in small batches[2]. (In the previous specification, the influence of these two variables would be reflected by ϵ_{it} and, perhaps, S_{it} if, say, large firms tended to have better educated managers.) If E_{it} and X_{it} were observable and had roughly lognormal distributions across firms,[3] their influence on the aggregate diffusion growth curve could be derived quite easily. The main empirical advantage of including such variables in the analysis would be that a more informed explanation would be possible of σ_j^2, the variance of the error term in (5.2.7.) As can be seen from the results of chapter 8, the cross-industry explanations of $(\widehat{\sigma/\beta})_j$ have had only moderate success and such an extension would seem particularly worthwhile.

1. Chapter 5, sections 1 and 2.
2. Which are particularly advantageous for the use of NC (see Appendix one,)
3. See pp.18-9 of Appendix one for evidence that X_{it} in this case may indeed have a positively skewed distribution similar to the lognormal.

Unfortunately, it would be well nigh impossible to obtain measurements of E_{it} for all firms in all industries and although, in this particular case, X_{it} could probably be measured, this would not help the cross-industry explanations of $(\widehat{\sigma/\beta})_j$. Batch-size of output is a relevant characteristic for three of the innovations only; it is true that technical factors could be identified for all of the innovations, but this would produce nearly as many explanatory variables as there are observations. Thus these technical variables must form an important part in individual case studies but their inclusion in a study such as this is, unfortunately, valueless. This does mean, of course, that for some of the structural parameters of the model (mainly σ_j^2), fairly large unexplained residuals will persist. To this extent, Gold et al's criticisms of the general model of diffusion are partly justified.

A second area in which more emphasis would be desirable is that of supply considerations. Unlike most past studies, some account has been taken of the role of the firms who supply the innovations, but this has been, necessarily, at a fairly ad-hoc level. The relaxation of two implicit assumptions of the model would suggest a more fundamental role for supply considerations, however. By and large it has been assumed (a) that the pricing of the new process is purely dictated by costs and (b) that supply is not constrained by capacity shortages.[1] Unfortunately, as stated earlier,[2] there is not sufficient evidence to assess how inappropriate these assumptions are: no evidence has been found that suggests that capacity constraints <u>were</u> important for any of the sample innovations, but this is by no means conclusive. Two alternative assumptions, which are no more or less plausible than (a) and (b) above, are that suppliers (c) price using a mark-up which is increased in times of booms and decreased in times of troughs and (d) increase their capacity in line with the <u>underlying</u> growth in demand for their innovation. Now, the model predicts that the demand for new innovations

1. See pp. 31-3 of chapter 4 for a brief examination of these assumptions.
2. Chapter 4, p. 27.

grows along an underlying bell shape[1] with cyclical variations superimposed on top of the trend. The implication of assumption (c) is that demand will not cycle as widely as in the constant mark-up case, since increased price will choke off some demand at the peak of the cycle, and decreased price will encourage extra demand in times of trough. Secondly, following assumption (d), the cycles in desired demand will not produce cycles in observed demand until after the peak of the bell shape is attained, thereafter, of course, excess capacity will exist, and supply will react more passively to changes in demand. Initially however, the cycles in demand will merely be reflected in lengthening and shortening of order books. The upshot of this is that the cycles in diffusion may be very mild or even non-existent for the first part of the diffusion period and will only appear with any strength in the second part of the period. This may be the explanation of the general non-significance of cyclical variations in the time-series curve-fitting of chapter 6.[2] At any event, some empirical evidence on these alternative hypotheses would be particularly helpful to the further extension of the theory of diffusion.

3. Areas for future research.

There seem to be two alternative directions in which future research on diffusion might progress. First, as already mentioned in this chapter, there are certain aspects of the typical diffusion process which remain partly uncharted, e.g. the behaviour of innovation suppliers, the role of certain technical factors. Moreover, the whole area of intra-firm diffusion has been largely ignored by this and most other contributions in this field.[3]

1. Which will be skewed, of course, for lognormal diffusion. Note that although the cumulative number of adopters grows along an S-shape, the number of new adopters (and thus actual demand, abstracting from repeat buyers) will grow along a curve which is the first difference of the S-shape, i.e. a bell shape.

2. For many of the sample, diffusion had not progressed into the second stage of falling demand at the time of data collection.

3. For a partial explanation see chapter 1, p.1.3.

Second, and even more marked by its absence in past work, is the study of the broader implications of diffusion for other areas of economics. Nowhere is this more marked and unjustified than in the theoretical and empirical study of growth economics.

To a certain extent, the theory of economic growth may continue to ignore diffusion so long as it remains pre-occupied with steady-state growth paths. In a recent comment[1], Matthews likens the rate of diffusion to the savings rate: it does not influence the rate of (steady-state) growth, even although it does determine the level of output and productivity. Although Matthews provides no formal proof of his assertion, it is undoubtedly correct and within the spirit of steady-state growth economics, as the following simple model illustrates.

Aggregate labour productivity is the weighted sum of labour productivity on all existing vintages of equipment:

$$Y_t = \sum_{v=T_t}^{t} w_{vt}\, y_v \qquad (10.3.1.)$$

where y_v is the labour productivity of vintage v equipment, w_{vt} is the proportion of the labour force working vintage v equipment, T_t is the date of construction of the oldest existing equipment and Y_t is aggregate labour productivity, all at time t.

Let there be two components to technical progress, (a) a constant, continuous improvement in best practice labour productivity at the rate g:

$$\frac{y_{t+1} - y_t}{y_t} = g \qquad (10.3.2.)$$

and (b) a constant diffusion rate of ϕ:

$$Y_{t+1} = (1-\phi)\, Y_t + \phi y_{t+1} \qquad (10.3.3.)$$

where ϕ % of the total labour force is switched from old machines to new vintage equipment every time period.

1. R.C.O. Matthews in Williams, (1973, op.cit.), p286

Combining (10.3.2.) and (10.3.3.) provides an expression for the growth in overall productivity of

$$\frac{Y_{t+1} - Y_t}{Y_t} = \phi\left(\frac{y_{t+1}}{Y_t} - 1\right) = \phi\left\{(g+1)\frac{y_t}{Y_t} - 1\right\} \qquad (10.3.4)$$

Solving (10.3.4) for the steady state growth rate h (say),

$$h = \phi\left\{(g+1)\left(\frac{y}{Y}\right)_t - 1\right\} \qquad (10.3.5)$$

Therefore, $\left(\frac{h}{\phi} + 1\right)\frac{1}{g+1} = \frac{y}{Y}_t$ (10.3.6)

Thus, in the steady state, the ratio of best practice productivity to aggregate productivity is constant. As y grows at the rate g, then so too must Y. Therefore $h = g$ and the diffusion rate determines only the level of Y, not its growth rate. This is, of course, analogous to the role of the savings rate, s, in standard Harrod- Domar or neo-classical growth models.

Furthermore, the stability of this growth path can be shown. For instance, if the diffusion rate were to increase to $\phi + \Delta\phi$ it can be shown that Y_t will grow at a rate in excess of g for a number of years, but once the new equilibrium ratio for (y/Y) is reached:

$$\frac{y}{Y} = \left(\frac{g}{\phi + \Delta\phi} + 1\right)\frac{1}{g+1}, \qquad (10.3.7)$$

growth in productivity will settle down, again, to the rate g.

Needless to say, such a model is at a high level of abstraction and is open to a number of criticisms. Are there any reasons why the best practice productivity should increase at a constant growth rate, or why the diffusion rate should be constant? One might expect that the rate g, itself, might be influenced by the diffusion rate, ϕ. It seems highly probable that where existing techniques are diffusing rapidly, the rate of invention and of innovation may also be faster due to technological feed-back. For instance, the problems encountered in the use of a new process often stimulate modifications to existing equipment or new inventions in adjacent stages of the production process. However, even abstracting from these other considerations, the fact that the diffusion rate does influence the <u>level</u> of productivity, and the rate of growth in periods of dis-equilibrium, is sufficient reason for the attention of empirical research.

Having said this, it must be admitted that most empirical studies of economic growth have tended to overlook the role of diffusion. This is not altogether surprising as serious problems of specification arise in the usual methodologies used in this area.[1]

Nevertheless, there are a number of peripheral questions which might be explored with little difficulty. For example, it would be interesting to assess whether the so-called 'technology gap'[2] between the U.S. and the U.K. is due to sluggish innovative performance or to characteristically slower diffusion of innovations, or both.

The following example provides one instance of how results, such as those presented in this thesis, might be used alongside such a study. Suppose data was collected on the date of innovation and diffusion, for a sample of innovations in both countries. As a measure of the technology gap, consider the difference in the use of these processes, say 10 years after their initial introduction in the U.S.

Assuming each new process had a simple cumulative normal diffusion curve in both countries, estimated by:

$$\hat{z}_{ijt} = \hat{a}_{ij} + \hat{b}_{ij}\, t_{ij} \qquad (10.3.8)$$

where z_{ijt} is the normit of the level of diffusion (Q_{ijt}) of innovation j, in country i in year t, and t_{ij} is the number of years after the first introduction of j in country i.

Suppose, further, that for innovation j, the U.S. innovated λ_j years before the U.K., then it would be possible to calculate the extent of the technology gap after 10 years and, more important, to assess the alternative ways in which the gap could have been avoided.

1. Such as the residual approach, as pioneered by E. Denison, See, for instance, "Why growth rates differ," Brookings Institution, Washington, 1967.
2. See the series of O.EC.D. case study reports: "Gaps in technology," O.E.C.D., Paris, 1968 and K.Pavitt, "The conditions for success in technological innovation," O.E.C.D., Paris, 1971.

As an example, let $\hat{\lambda}_j = 2$, $\hat{a}_{U.K.,j} = \hat{a}_{U.S.,j} = -1.6$ [1], $\hat{b}_{U.K.,j} = .16$ [2], $\hat{b}_{U.S.,j} = .18$

in which case,

$$\hat{z}_{U.S.,j,10.} = -1.6 + 1.8 = .2, \text{ therefore, } \hat{Q}_{U.S.,j,10.} = .579$$

Similarly, $\hat{z}_{U.K.,j.8.} = -1.6 + 1.44 = -.16$, therefore, $\hat{Q}_{U.K.,j.8.} = .374$

Thus, the gap is .205, i.e. 20% fewer U.K. firms have adopted as compared to U.S.

Had the U.K. innovated at the same time as the U.S. (i.e. $\lambda_j = 0$), then the gap would have been reduced to: .579 −.5 =.079. Alternatively, had the U.K. diffused as quickly, (i.e. $b_{U.K.,j} = 1.8$) the gap would have been reduced to: .579 −.436 =.143.

The gap could have been removed altogether by the U.K.

(a) innovating 1.25 years earlier than the U.S., i.e. $\lambda_j = -1.25$, or

(b) diffusing 40% quicker than it did, i.e. $b_{U.K.,j} = .225$.

Now, from the findings of this thesis, it is possible to explore the alternative ways of effecting (b) and, given a complementary study on the determinants of innovative performance, it might be possible to examine the alternative ways of effecting (a). In principle then, a cost-benefit analysis of the optimal policy mix in this context should be possible.

Conclusions.

This thesis constitutes the first broad-based inter-industry analysis of diffusion in the U.K. It has confirmed a number of previous American results, but also some new findings have emerged. Thus, further support is provided for the hypothesis that diffusion, typically, follows some sort of S-shape, the slope of which is at least partly determined by the profitability of the new innovation in question. Similarly, it does appear that large firms adopt,

1. Given that a has little economic meaning in the cumulative normal, this assumption seems fairly harmless. Its relaxation would merely complicate the example; unless widely different values are assumed for the two countries, the substance of the ensuing conclusions remains largely unchanged.

2. The average for the sample innovations in the cumulative normal acceptable set (table 6.4.1.)

on average, more quickly than small firms. In addition, however, evidence has been presented to suggest that the exact shape of the S-curve may be influenced by the type of innovation and, to a certain extent, the trade cycle. Moreover, the parameters of the diffusion curve appear to be sensitive to the structure of the industry concerned, albeit in a rather complicated way.

On the theoretical side, the model used is less mechanistic than the conventional epidemic model and, consequently, it is somewhat richer in the number of predictions it generates. Moreover, the quality of those predictions seems to be superior, at least for the data observed in this case.

Having said this, further research (and data collection) is still required on certain aspects of the diffusion process, notably, the role of the innovation suppliers. Similarly, the broader implications of results, such as these, require some attention if future empirical studies of economic growth are to be useful to policy makers.

Appendix I: Technical aspects of the innovations and industries studied.

This Appendix considers two important characteristics of the innovations and industries for which data has been collected. First, how representative are both of the overall U.K. economy and, second, can the innovations as a group be described adequately by a number of stylised attributes?

1. Representativeness of the sample

It should be stressed, initially, that the selection of innovations to be studied was in no sense random. Quite simply the sample constitutes 22 process innovations for which sufficient data has already been collected by other economists or is available in reasonably accessible published sources[1] (e.g. trade or scientific journals.)

Whilst it is considered that such a sample size is probably sufficient to support the econometric work described in the main text, it can hardly be claimed to be large enough to be representative of all innovations that are made in British industry, especially given its non-random selection.

Having said this, it is important to assess whether there are any obvious deficiencies in representation. Using Pratten's[2] classification of industries into four types: 'process', 'engineering and allied', 'textile and clothing', and 'others', table A.1.1. shows how the thirteen industries are distributed (some industries studied accounted for more than one innovation.) Clearly, overall, roughly ten per cent of all industries are covered with roughly the same percentage of all employees; at the slightly less aggregate level, the ten per cent figure is maintained for 'process' and 'textile' industries but exceeded slightly for 'engineering', with a compensating short fall for 'others'. Against this rather pleasing picture, it should be noted that within 'process', crude steelmaking accounts for three innovations; within 'textiles', weaving for four and within 'others', paper and board for five - these three industries then, are over-represented.

1. For more details see the following appendix.

2. Pratten, (1971), op.cit., p265.

Turning to the innovations, seven are within the 'others' category, with five each in the three remaining categories, thus maintaining a roughly representative spread.

It is probably impossible to say much about how typical are the innovations studied - to do so would require massive data collection and encounter serious definitional problems. Certainly it is probable that they are, on the whole, what might be termed 'major' innovations almost as a corollary of the fact that data has been collected about their diffusion. Table A1.2. is an attempt to classify them into broad functional categories. Only eight fit into the crude vintage model of replacement of an old technology embodied in existing capital equipment (K); six either automate old manual operations or do away with them entirely (L); six more 'supplementary innovations' can be applied to old existing equipment as well as to new equipment (S); and two perform functions that previously were outside the capabilities of existing equipment [1] (N).

Table A.1.1. The coverage of the sample.

	Total U.K. (1968)		This sample	
Type of Industry	Number of Census Trades [2]	Number of Employees(000) [2]	Trades	Employees (000)
Process	32	1376	3	260
Engineering	35	2938	5	347
Textiles	25	1163	2	106
Others	49	2601	3	104
	141	8078	13	817

1. It is interesting to compare this breakdown with that given by Mansfield for the 12 innovations within his sample (Mansfield(1968),op.cit. footnote 25 p.146): for the same categories he has 8k, 1L, 2S and 1N.

2. Source : C. Pratten, op.cit.

Table A.1.2. Classification of the Innovations.

Innovation	Industry	Type of Innovation*	Abbreviation used
Special Presses	Paper and Board	S	SP
Foils	" "	S	F
Synthetic fabrics	" "	S	SF
Wet Suction Boxes	" "	S	WSB
Paper machine control by computer	" "	L	PCBC
Gibberellic Acid	Malting	S	GA
Computer typesetting	Provincial Evening Newspapers	L	CT
Shuttleless looms	Textile weaving	K	SL
Electric hygrometer	" "	L	EH
Accelerated drying hoods	" "	S	ADH
Automatic size boxes	" "	L	ASB
Tufted carpet machines	Carpet manufacture	K	TC
Automatic track lines	Car manufacture	L	ATL
New methods of steel plate cutting	Shipbuilding	L	SPC
Numerically controlled machine tools	Printing press manufacture	K	NCPP
Numerically controlled machine tools	Turning machine manufacture	K	NCTURN
Numerically controlled machine tools	Turbines manufacture	K	NCTURB
Tunnel Kilns	Brick making	K	TK
Basic Oxygen Process	Iron and Steel	K	BOP
Continuous Casting	" "	K	CC
Vacuum degassing	" "	N	VD
Vacuum melting	Certain special steels	N	VM

*Here S implies that the innovation is a supplement or addition to existing plant, K implies that the innovation replaces an old technology embodied in capital equipment, N implies that the innovation serves a different purpose

(Chart A.1.2. notes continued)

than does any existing equipment.

L implies that the innovation involves automation.

Needless to say, this classification is somewhat arbitrary and there is, in practice, a certain amount of overlapping.

2. Characteristics of the innovations.

In order that the model building of this thesis should be applicable to the innovations within the sample, it is essential to consider the technical aspects of each innovation in turn. For the sake of brevity, each description is split into three main sections: a) A simple technical description of the innovation, its role within the overall production process of the industry concerned, its main advantages over the old technology and two major economic characteristics - cost and profitability.

b) A brief history of its invention, development etc., and, where possible, some reference to the industry which is responsible for its production and marketing.

c) A more detailed examination of the economics of adoption of the innovation with particular reference to the following possibilities:

1. Economies of scale in its application and, more generally, a consideration of any technical factors which might lead to it being more profitable for larger firms.
2. Improvement over time in its capabilities (including price).
3. Non-linearities in its short run average cost curve (i.e. sensitivity of cost savings to capacity usage).
4. Changes in the quality of the end product as compared with the old technology.
5. Problems of co-ordination with the existing equipment used on adjacent stages of the production process.
6. The effects (if any) on the profitability from adoption of differences

between firms in their products, processes and inputs. This leads on to an assessment of the technically feasible set of potential adopters.

Wherever possible, references have been made to more technical discussions of the innovations and, due to the preponderance of footnotes, they have been grouped at the end of this appendix.

(i) Special Presses (SP) in Papermaking

a.) The paper machine represents the major technical stage in the process of making paper and board from the raw material of woodpulp or waste paper: it transforms the web (a dispersion of fibres in water) into paper sheet. The paper machine itself consists of three sections: the wire section, the press section and the drier, one role of each of these being to reduce the water content of the web. Special Presses, which operate at the second stage, press the water from the web and remove it, either in cavities in the rolls of the press, or by means of special fabrics. [1] The water removal performance of SP is much better than that of the old technologies (solid and suction presses), and it enables the paper machine to be run at a much faster speed. In addition, the quality of the end-product is often improved.

The investment cost will depend on the width and age of the paper machine, ranging from £8,000 to £40,000 [2]; the average cost given by adopters, when questioned in connection with the N.I.E.S.R. study, was £14,700. The profitability of installation will also vary depending on the same factors; one estimate of the typical pay-back for the U.K. industry is given as 1.6 years, [3] and the average estimate of firms in the U.K. N.I.E.S.R. sample was even lower - only 1 year.

The vintage approach seems inapplicable in this case, as SP can be installed on new or old paper machines, although they may be more profitable on the former.

b.) The early work on invention was motivated by discontent with the existing technology, and although the early patents in the late fifties were Swedish, the first commercial introduction was in the U.S. in 1962. In the U.K. the first adoption was in 1964 and there are, at present, four machine makers offering their own alternative Special presses. Three of these are subsidiaries of American firms and one is a subsidiary of a Swiss firm.

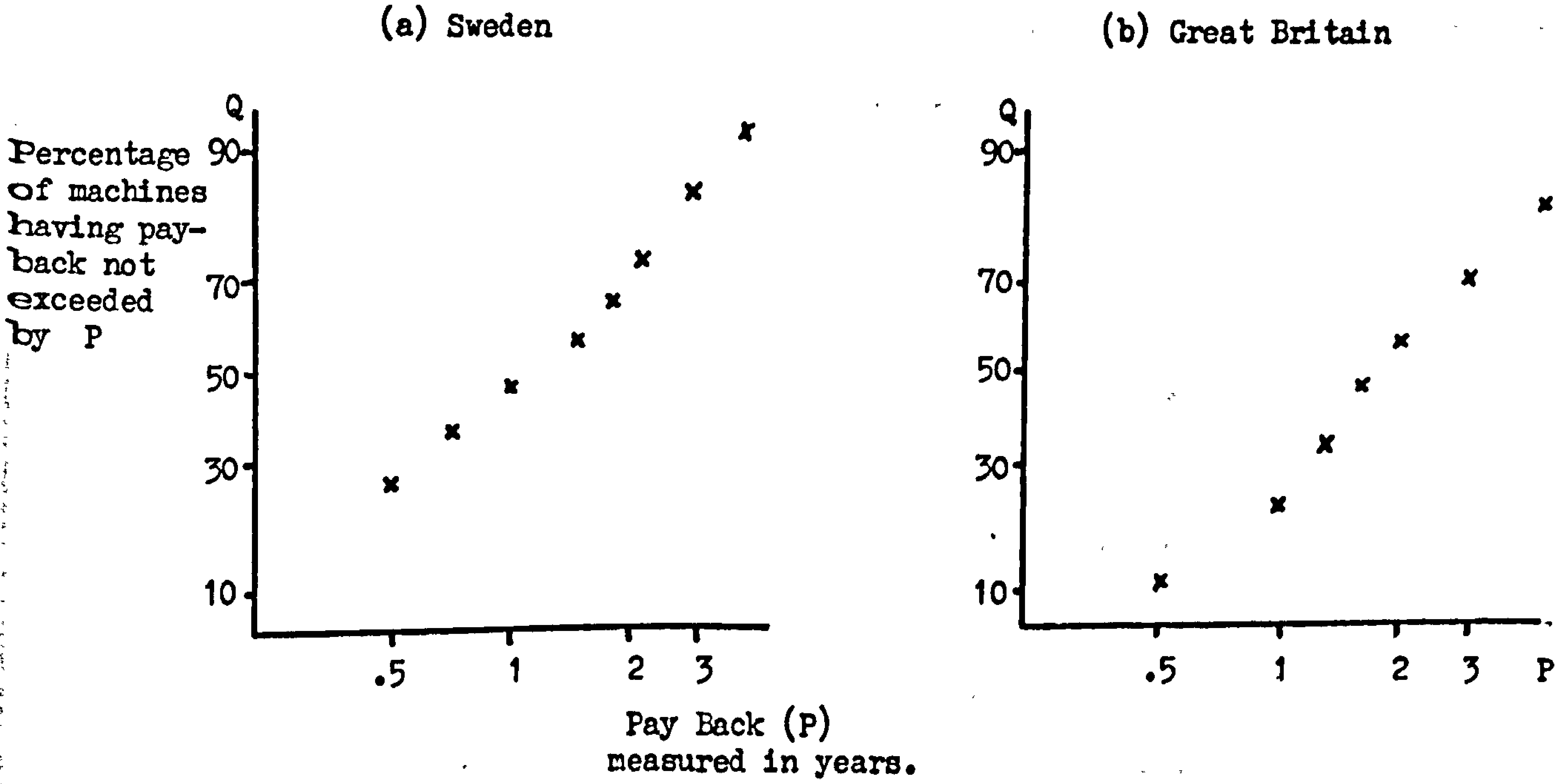
c.) Larger firms might find the innovations more profitable in that they are more likely to own paper machines for which SP are best suited ('new and narrow'). This might be the case i) simply because larger firms own more machines of all types and ii) perhaps their capital stock is likely to be newer, on average, than that of small firms. This is generally borne out by data presented by Hakonson[4]; he concludes that 'the pay-off period .... appear(s) to be related to size.'[5]

As Hakonson points out [6], however, the profitability of installation depends crucially on the adopter being able to sell the increased output and so one might expect profitability to vary over the trade cycle.

Post-invention improvements in the technology have resulted from each producer bringing his own version of the innovation on to the market with apparent advantages over his competitors' products, this, in turn, provoking a reaction from the others in an effort to improve their product.[7]

The feasible set of adopters has been defined as all Paper and Board manufacturers. Quite clearly, profitability of adoption will vary across these firms depending on their existing capital stock and the nature of their product. Hakonson calculated hypothetical pay-back rates (based on actual technical data collected from firms) for 107 machines in Sweden and 146 in the U.K. [8] Using this data to construct cumulative frequency distributions of pay-backs, one finds that, in both countries, a non-symmetrical S shape is plotted for increasing values of pay-back. Indeed, when these cumulative frequency distributions are plotted on log probability paper, against the pay-back period, a straight line gives a remarkably good fit in both cases. (see Figure A.1.1.) This would suggest that pay-back may be lognormally distributed across machines in these industries.[9]

Figure A.1.1. Profitability distributions for Special Presses.



Source: Hakonson, New Industrial Processes p.92.

(ii) Foils (F) in Papermaking

a) Very crudely, the main aim of this innovation (as is true also for Wet Suction Boxes and Synthetic forming fabrics) is the removal of moisture at the press stage of paper production. The advantage over the old method - grooved table rolls - is that the water is removed in a more orderly and controlled fashion, which leads to possible increases in machine speeds (and thus output) and improvements in quality (reduced wire marks, two-sidedness and fewer pinholes.) The average cost of a Foil Installation (including Forming Board, Foil Unit and Drainage Control Unit) in 1974 was £2,500 [1] (this applies to only one paper machine.) Seen against the average size of firm in the industry, then, this is a relatively cheap innovation. Factor savings are mainly indirect, in that with the same inputs, output is increased (physically and perhaps price-wise due to improved quality.) A typical estimate of the pay-back period is about 6 months.[2]

b) The invention appears to have originated in North America, early development work being carried out in the late fifties by Wrist and Burkhart in Canada and Lodding Engineering in the U.S.A. The innovation is made available by at least three firms in the Paper machinery industry in this country, which is notable i) for its geographical concentration in or near Bury and ii) for the fact that virtually all its firms are foreign owned [3].

c) A serious limitation to the profitability of installing foils on existing paper machines would be deficient demand; there is very little point in speeding up output rates unless the saved time is also used for extra production. The stage of the trade cycle might therefore be an important determinant of the immediate rate of return to be gained from adoption.

There have certainly been improvements in the performance of foils over time. Technical improvements were made that removed quality impediments and decreased physical wear and tear; but perhaps more interesting are the improvements in applicability of foils. Initially, they showed little advantage over table rolls at speeds (of the paper machine) below 200 feet per minute; however, continuing development produced foils that were not only advantageous for all speeds but which increased the savings achieved on faster machines as well.

The manufacturers claim that foils are easily adaptable to any existing paper machine and that problems of non-compatibility are unimportant; typically, they help install the innovation.

It is difficult to see why any paper maker should not be able to use foils, although profitability in their use will to a certain extent be governed by the type of paper produced and possibly the age of the paper machine.

The feasible set of firms may be defined, therefore, as all Paper and Board makers as classified by the British paper and board manufacturers Association.

(iii) Synthetic wires fabrics (SF) in papermaking

a) Again this innovation is relevant to the removal of water in the production of paper; it is reckoned to be superior to the old technology (bronze wires) in that it yields superior drainage and increases the speed of the paper machine (due, partly, to greater life expectancy and durability resulting in large reductions in down-time for the paper machines.) The investment cost is roughly three times that of bronze wires - about £600 in 1971.[1] On the other hand, the life expectancy might be at least four times as long and a pay back of 3 months might reflect the cost advantages of synthetic wires. Clearly the vintage approach is inapplicable - a new paper machine is not necessary in order to use synthetic wires.

b) It has proved difficult to trace the history of this invention which dated from the turn of the sixties, although certainly Porritts and Spencer of Bury have played an important part in developing and marketing. Once more this is a firm outside the Paper and Board industry.

c) Until very recently, synthetic wires have been most advantageous on smaller machines (as the life expectancy differential over bronze wires is greater), but this should not necessarily be taken to mean that smaller firms will find the innovation more profitable; certainly smaller firms will tend to own small machines, but large firms may do also, especially given the fact that the paper grades typically made on smaller machines are by no means the speciality of small firms. All that can be said is that small firms should not be at a serious disadvantage. The same sort of considerations regarding the effect of the trade cycle should operate as was postulated for foils and SP as, again, the main advantages of the innovation lie in increased output from much the same inputs.

Recent developments have made the innovations more advantageous than previously for larger machines. An interesting point to emerge from communication with the manufacturer is that 'each time a firm makes its first trials with synthetic wires, there can frequently be a history of failures or only partial success, the cost of which is usually borne by the wire manufacturer.' To the extent that this is true, much of the risk of adoption is removed for the prospective adopter; on the other hand, as is pointed out in the same letter, the consumer pays for this 'development work' in the form of a higher price of the innovation.

Exactly the same remarks about the feasible set of potential adopters are true for synthetic wires as were true for foils and SP; again, potential profitability will depend partially on each firm's product and paper machine.

(iv) Wet suction boxes (WSB) in papermaking

a) Of all the innovations in this sample, this has proved to be the most difficult to document. Yet again, this is considered to yield superior drainage from the paper machine and results in increased production and improved quality as compared with the old technology - again table rolls; however, wet suction boxes are not altogether a substitute for either rolls or foils - manufacturers suggest that they should be used in combination. Having said this, it is probable that the addition of wet suction boxes to a machine which is already using foils will be less profitable. Compared with a machine using only table rolls, the speed of the machine can be increased by as much as 25%. A typical installation of Wet suction boxes might cost £3,000 with a pay-back period of 3 - 5 years, although this will vary widely depending on circumstances. [1] As with the three previous innovations, WSB may be termed a supplementary innovation which does not involve replacement of existing paper machines.

b) The innovation was developed as a co-operative venture between equipment makers (K.M.W.) and the Control Paper Laboratory in Stockholm, Sweden in the 1950's. [2] In this country it is marketed by two major paper machinery makers and would account for a very small percentage of their total revenue. Slightly differently from foils, most of the advertising literature concentrates on the advantages of Wet suction boxes, as opposed to other methods of de-watering, rather than emphasising advantages over the other 'brand.' Possibly this is due to the innovation's much slower take-off (see the following Appendix, section 4) and relatively lower profitability.

c) Although it is claimed that 'higher production and improvements in quality' can be obtained on almost all machines, it is apparent that higher speed machines are better suited, and, to the extent that larger firms tend to employ faster machines, they have most to gain from adoption. Once more, for reasons already stated, the trade cycle will influence the profitability of adoption.

Again, the innovation should be applicable to all paper makers, although its profitability will vary depending on age, width and speed of machine.

(v) Process Control of Paper Machines by Computer (PCBC) in paper making

a) Process control by computer is, of course, an extreme case of automation; in this case the computer has control over a number of variables in the paper making process - machine speed, thick stock flow, clay flow, flowbox pressure and slice gap etc., or, in layman's English, it controls the speed of the whole process, the rate at which various materials are added to the basic wood pulp and the speed of various sub-sections of the process. Consequently, the operator should have greater and more immediate control over his machine, leading to increased machine speeds without any loss in quality; downtime is also reduced for a number of reasons e.g. switching from making one grade of paper to another can be effected very quickly. It is also claimed that, as paper quality is much more consistent, waste due to so-called 'off-specification paper' is substantially reduced.

As opposed to the other innovations in paper making considered, process control represents a major investment for all paper firms, the cost of a relatively small installation in 1964 was £100,000. On the other hand, returns can be substantial - a typical pay-back might be two years [1] and this would be due basically to increase in output from fixed inputs, or perhaps even reduced labour and material inputs.

Again the vintage approach is not altogether appropriate as process control can theoretically be applied to any vintage paper machine but would probably be less profitable for very old vintages. Moreover, no old technology equipment is replaced.

b) The history of computers in general and process control in particular is well documented elsewhere, [2] suffice it to say here that the concept of process control was not introduced initially particularly for paper making. The first computer firm to sell a process control computer in the Paper industry was Elliott-Automation in 1964 [3]. To date, there are four firms in the computer industry that have sold process control to the Paper industry (none of them with any financial interest in the latter industry).

c) The extent of returns to be gained from adoption will depend on the size of the paper machine and thus of the computer. Generally speaking, the capital cost of a computer does not rise proportionately with its size.[4] Further, certain specialised training is necessary for senior work staff, and large firms may be more able to release them for such training. Similarly, a firm running a large number of paper machines may be more able to cover the loss of output in the change-over period. Reference is made by a machine maker [5] to a period of six months, after which 'the computer and control installation had been very successfully merged and adopted into the overall working routine.' Potentially, then, certain liquidity problems might arise during the adoption period.

Some of the gains resulting from adoption are probably independent of the level of demand for paper (notably quality improvements) but others (increased running speeds, for instance) will be of less importance during the downturn of the trade cycle. As computers as a whole have undergone a period of fast technical improvement, it is probable that the real costs of installation have decreased and the relative advantages of the technique have increased over the period.

The feasible set has been defined as all Paper and Board makers, although the high installation costs present a serious barrier for small firms.

(vi) Gibberellic Acid (GA) in Malting

a) The Malting industry transforms grain (usually barley) into malt, most of which is then used by the brewing industry (although many brewers actually do their own malting.) It has three main processes: Steeping, that is soaking the grain in water for two to three days; germination in which the drained barley loses weight and is transformed from a hard mass into the softer malt - this can take from one to one and a half weeks, and finally, drying and curing, in which the malt has its excess moisture removed in a kiln (often for up to twenty-four hours.) Gibberellic acid is an additive which is usually applied at the germination stage but sometimes at the end of steeping. Its main effect is to reduce germination time to as little as five or six days; other side advantages claimed are reduction of malting loss, increases in the yield and quality of the extract and reduction of the dormancy period (a period in the growth cycle of barley which must have occurred for germination to take place.) [1] Although the adoption of GA in itself incurs no capital cost, to use it efficiently does require the purchase of specialised machinery (a moving spray boom, a reservoir for the GA solution and pipelines); in 1967 this would have cost approximately £500. The pay-back period attached to adoption varies greatly depending on local circumstances (see below) - seven firms who answered a question on this from the N.I.E.S.R. gave an average answer of 1.5 years. Figures given in a Trade Journal [2] for one actual installation of GA and appropriate equipment, tend to suggest a much shorter pay-back period, with labour productivity rising by over thirty per cent in the first year. However, in this case, there had been overall improvements in the plant and it might be fallacious to attribute the total increase in productivity to the adoption of GA.

b) The innovation probably dates back to basic research carried out by a Japanese scientist in the 1920's. After the war in Europe, pure GA was first extracted by a team of I.C.I. scientists, although the first public report on it was by a Swede to the European Brewing Convention. I.C.I. is the leading producer in the world of GA and indeed, in this country, it is the only producer. This can be claimed as one of the few British innovations in the sample, but again its producer is not connected directly with the consuming industry. In promoting GA, I.C.I., as a first move, informed all maltsters and brewer-maltsters of the innovation and its capacities.

c) There are no inherent attributes of GA that should lead to scale economies in its application. Nevertheless, the returns yielded do depend on the type of grain used, the age of existing capital equipment and, to a lesser extent, the type of beer produced; larger firms, just because of their larger scale, must have a higher probability of having the optimal conditions for adoption. Moreover, Ray [3] stresses that a certain amount of sophistication is required of managers in the way that GA is administered; he hypothesises that "(smaller) companies usually cannot afford to have professional managers or scientists who can provide the sophistication required" and thus, if this is the case, GA may be less profitable for smaller firms.

The main improvements over time in the technology have been i) a reduction in price of GA itself (falling from £.875 per gram in 1962 to £.58 in 1974),[4] presumably due to learning by doing by I.C.I. and ii) better methods of applying GA - better, that is, in the sense of more efficient and uniform. One might put this down to learning by doing by the malsters.

The state of overall demand for the product is considered to be important in assessing potential gains; given the largely fixed nature of the general capital equipment and, to a certain extent, skilled labour in a malting plant, GA cannot so much be used to reduce inputs needed for a given output, as to increase output from a given stock of inputs. Clearly, this is only of use if the extra output can be sold.

The feasible set of firms has been taken to be all malsters and brewers which perform their own malting - that is, brewer-malsters. Within this set, however, potential profitability from adoption may vary considerably. Most notably, the age of the kiln used can reduce potential returns; often older kilns operate as an effective bottleneck [5] - which can only be removed by replacing the kiln - in this case, there are only limited returns to be gained from speeding up an earlier stage of the production process through the use of GA. Secondly, some malsters may be limited in the extent to which they can use GA by some of their customers refusing to accept GA-treated malt. [6] Finally, the effectiveness of GA in speeding up germination will depend on the characteristics of the barley used.[7]

(vii) <u>Numerical Control of Machine Tools (NC) in Three Branches of Engineering</u>

a) Numerically controlled machine tools are simply machine tools which are controlled and monitored in their operation by a computer of some sort. A more rigorous but perhaps less comprehensible definition might run as follows: "Numerically controlled machine tools are controlled by the input of numerical information carried on a punched card or magnetic tape and transmitted by a programming office; electronic adapters attached to the machines convey the information to them as control commands." Because continually changing processes can be handled on the same machine, this allows <u>flexible</u> automation of the old technology.

The main economic advantages are: labour savings, as often one numerically controlled machine tool can do the job of two or three conventional machines; savings on other tools, such as templates, jigs and fixtures; greater accuracy and uniformity of end-product i.e. quality improvement; more flexibility, allowing lot sizes to be tailored more closely to requirements and therefore reducing the need for stocks of finished products; and finally, NCMT may sometimes be able to produce part designs that are impracticable using conventional machine tools.[1]

The cost of NCMT varies widely depending on its type: in 1968, simple drilling machines varied between £3,000 and £18,000; milling machines between £6,000 and £75,000 and machining centres between £20,000 and £250,000 [2]. The average price paid by firms in the N.I.E.S.R. sample was £20,700. Consequently, one might also expect the profitability of installation to vary widely: in another N.I.E.S.R. survey [3] the average pay-back quoted by thirty-seven firms was 5.41 years, and Mansfield quotes an average pay-back given by twelve firms in the American tool and die industry as 4.8 years.[4]

This is a relatively 'unlumpy' innovation, as one NCMT can be installed without replacing a large proportion of the existing stock of machine tools. The vintage approach is reasonably appropriate.

b) The concept of numerical control probably first suggested itself to the American aircraft industry when looking for a faster and more accurate method of producing complicated parts and components after the war.[5] Development work was started in that country in 1947 jointly by the industry, the Air Force and the Universities; a prototype was presented in 1952 and the first industrial application took place in 1955. Most of the early development work in Europe took place in the U.K. and the first industrial application followed in 1956.

There are probably over a dozen firms in the British Metal working machine tool sector that are currently producing NCMT and in addition, imported NCMT take a large share of the market. Based simply on sheer numbers, one might conclude that this is the most competitive supplying industry of any of those considered in this sample; however, it should be mentioned that certain firms specialise in certain types of NCMT. In the case of the diffusion of NC in the Turning machine tool industry, (NC TURN) the supplying industry coincides with the consuming industry, although it should be stressed that many producers of turning machine tools do not produce NCMT.

c) The profitability of installing NCMT depends on the nature of the end product and the typical batch size. As large firms will tend to produce a range of products, they must have a higher probability of producing under the conditions for which NC is most suitable. Further, NC requires sophisticated management and expert staff and it might be that these are more readily found in larger firms. The relatively high cost of NC might also work against early adoption by small firms, Gebhardt reports[6] that the price of a single NCMT was higher than some firms' total average investment in a year. Further, Ray finds, in the U.K. at least, that 'economies of scale in the servicing operations lead to a higher rate of return being yielded by relatively large scale installations of NC machines.'[7] On the other hand, Gebhardt argues that smaller firms are often more suited to numerically controlled machines because they usually produce smaller batches than large firms.[8] This argument <u>may</u> be relevant when considering the overall profitability of NC to a firm's <u>total</u> production programme but as this is a study primarily of the imitation process, we are concerned only with the <u>first</u> adoption by any firm. Even if small batch products are <u>relatively</u> less important for large firms, they may still produce <u>absolutely</u> the same amount or even more of such products.

Similarly, one cannot be certain to what extent the technology of NC has improved over time; Gebhardt[9] talks rather vaguely about 'continuous technological improvement' and of certain firms refraining from purchasing because of the 'too rapid technical development of the control features.'[10] Ray, also, refers to the technology developing very rapidly.[11]

It is possible to derive an implicit price index from the Business Monitor series for NCMT for the years 1966 - 70. In that period, unit price rose by 113% - far in excess of the overall rate of inflation. Presumably the main influence on this massive increase in price was a similar increase in the quality of the machines, although it is likely that any learning by doing in the manufacture of NCMT had lessened by this time.

Given the heavy overheads incurred by NC (sometimes the firm installs its own computer), high capacity usage is essential, in order to derive the maximum returns from installation; on the other hand, to the extent that NC produces a higher quality product and even, in some cases a unique product, this might act as a buffer against short-term fluctuations in overall demand.

Because it embodies such a radical change in the production process, NC is bound to cause certain co-ordination problems; retraining of old staff, employment of new staff, re-arrangement of factories and trade union resistance [12] may all occur.

The diffusion of this innovation was studied in three small parts of the Mechanical Engineering Industry: by firms producing Printing Presses, Turbines and Turning Machines - in each case, the innovation was considered suitable for all firms. Reference has already been made to the typical batch size produced by a firm as an important influence on the profitability of adoption; Gebhardt [13] presents data from a German machine maker showing the relationship between production costs of hand-operated, numerically controlled and automatic machines and batch size and number of orders. Generally, NC has lowest costs for batch sizes ranging from five to fifty; however, the more often the batch production is repeated, the lower is the batch size for which NC becomes more profitable than hand operated machines; for instance, if the order is repeated ten times, NC becomes the most efficient method for batches varying between one and forty; if the order is a one-off, then NC is most efficient over the batch range 25 to 300. Another factor determining profitability is the type of machine - Gebhardt presents a table showing relative cost savings actually achieved by firms in his sample, from which it is clear that even for the same batch size, certain NC machines are more profitable than others in practice, depending on the nature of the end-product [14] (e.g. how sophisticated it is). In a later table [15] he shows the incidence of batch sizes in the firms studied; certainly in the U.K. (but in other countries also) the distribution is positively skewed i.e. relatively smaller batch sizes have a preponderance.

(viii) Shuttleless Looms (SL) in Textile Weaving

a) The shuttleless loom is certainly the most revolutionary recent innovation in the weaving stage of textile production. In the traditional process, [1] the cotton threads are woven by a shuttle (carrying a thread) moving to and fro across a series of threads arranged lengthwise. Shuttleless looms use other methods of transmitting the single thread across the lengthwise threads (or, technically, have a different method of weft insertion). The Sulzer[2] shuttleless loom, for example, draws the weft (thread) from a stationary supply package mounted on the side of the loom, instead of, as with the traditional method, carrying the supply with it. The major advantages derive essentially from i) the higher speeds possible and ii) less labour time being taken up in replacing the weft supplies - the packages used by the shuttleless looms having a far greater capacity than the shuttles used in the traditional method. This results in higher production from a given floor space and increased labour productivity, in addition to an improvement in the quality of the finished product (because of fewer weft breakages and less strain on the lengthwise threads).[3]

The cost of an average Sulzer loom (£3,500 in 1962) [4], whilst up to five times more than that of a conventional automatic, means that it is a relatively cheap innovation (although usually looms are installed in batches.) There are wildly varying estimates of the profitability of replacing a conventional with a shuttleless loom: a Textile Council Report (1969) suggests a rate of return of 22%; however, four firms actually using SL gave an average estimate [5] of ten per cent.[6]

This is clearly a new technology which can only be employed if latest vintage equipment is bought - it cannot be latched on to old equipment -

and is thus roughly in line with the vintage approach.

b) The history of the invention of the shuttleless loom is very long [7]; relevant patents were taken out in this country and abroad in the last century, but most of the important development work was carried out by a Swiss firm (Sulzer Bros.) after about 1930, (other types of shuttleless loom have been developed in various other countries since the war.) The first commercial introduction was probably in Germany in the early fifties and there are a variety of looms available to the British industry, all produced by foreign machine makers (from Switzerland, Spain, U.S.A., Japan and Czechoslovakia). Competition between machine makers is international but Sulzer Bros. are almost certainly the dominant firm.

c) Broadly speaking, the size of a loom is fixed and so any economies of large scale that exist must derive from being able to install larger numbers of shuttleless looms, to quote the N.I.E.S.R. study '(adoption of) small numbers of shuttleless looms would hardly be worthwhile in the case of Sulzer Looms' [8] and 'larger firms have greater opportunities to optimise the ratio of operatives to looms and can more easily achieve longer runs between stops for fabric changes, whereas small firms may not be in a position to organise multi-shift working.' [9] (Often, one operative may be operating twenty different looms, and it should be remembered that large firms may sometimes have a strong bargaining position vis-a-vis their workers in towns where there are few other employers.)

N.I.E.S.R. also note that 'the new looms are now of a performance significantly superior to those made five to ten years ago;' [10] in addition, they appear to have widened their applicability to cloths of increased width and multi-colours.

Cost savings depend crucially on long production runs and shift-working [11]; clearly, in that case, the state of demand will be of importance in determining the profitability of installation; at some low capacity usage rates, total costs of SL may well be in excess of the operating costs of the oldest conventional machines.

In theory, the shuttleless loom could be adopted by any weaver (defined as a firm in the textile industry), at least partly weaving cotton or man-made fibres.) In practice, abstracting from the influence of scale, there are a number of reasons why the profitability of installation will vary across firms. The lay-out of the weaving shed, if old-fashioned, presents problems - not only will adoption cause disruption and extra work for management but it will also probably yield lower returns, especially where there are physical limitations to the possible improvements in the old factory. Similarly the nature of the product may reduce profitability; as the Institute notes, 'relative savings are far smaller where the cloth being woven is too heavy or too fine' (only certain shuttleless looms are physically capable of coping with heavy yarns at all, and labour costs are proportionately less important when costly fine yarn is woven.) A reason that non-adopting firms often give for not having adopted is that their output is insufficiently standardised thus reducing the chance of the long production runs necessary to derive large savings from adoption.[12]

It would be wrong to conclude that firms with these special product characteristics are not potential adopters; given different demand conditions and sufficient investment, adoption could well prove still to be beneficial. Nevertheless, it is obvious that the potential costs of and returns from adoption must vary considerably across firms in the industry.

(ix) New methods of steel plate cutting (SPC) in Shipbuilding.[1]

a) Steel plates are used extensively in the construction of all new ships above a certain size. Traditionally, plate is bought from the steel industry and cut by hand by relatively skilled workers, after having previously marked the appropriate shapes in the plate, also by hand. The two new methods considered here - photo-electrically controlled cutting machines (Ph.E) and numerical control (NC) - both eliminate the marking process and automate the cutting process. In Ph.E, a drawing of the shape required is placed in the control side of an automatic burning machine, which then cuts the plate to the appropriate shape with the help of a photo-electric device. In NC, the negative of the drawing is transferred to a tape programming the cutting machine.

Obviously the main aim is to save direct labour costs[2], but other lesser advantages would arise in the form of more economical use of steel plate and increased precision. On the other hand, certain extra labour is needed - programmers and/or 'photographic staff'. The cost of a typical installation was £54,000 in 1957 prices, [3] which represents only a medium sized investment for most shipbuilding firms. One estimate of profitability is a payback period of ten years which is approximately confirmed by the estimates given to N.I.E.S.R. by adopters (average of just under twelve years payback.) [4]

b) It is difficult to attribute these new techniques to any unique source, serious development work was carried out by a number of machine makers in various countries in the late forties and fifties. In a communication from the British Ship Research Association, the main current manufacturers were listed as two German, two Norwegian and two British firms (B.O.C. and Hancocks). Clearly then, the shipbuilding industry is not responsible for the innovation itself, rather it is dependent on an international oligopoly.

c) Although there is no published information on economies of scale, one might suppose: i) that where NC machines are computer controlled, capital costs would not rise proportionately with the number of machines, ii) that to the extent that the fixed costs are high (including specialist staff, high installation cost meaning large interest and depreciation allowances), large shipyards may benefit from being able to use the machines more intensively.

Generally, NC is an improvement over Ph.E and both have been subject to improvements in scope (e.g. NC was first only applicable to plain and single curved plates but later, much more complicated exercises were possible - notably three dimensional surfaces.) One might say, therefore, that 'the innovation' has become increasingly more profitable (other things remaining equal) over time.

One major advantage of both variants claimed is that they remove a bottleneck (the traditional method sometimes holding up other stages of the production process) and thus increase effective capacity[5], but quite obviously in periods of low demand, the removal of this bottleneck will be less necessary and, consequently, returns from adoption in such periods will be that much lower.

This is certainly an innovation which potentially has disruptive aspects for both management and work force - not only are new skills necessary of the latter but the former will also need to employ new staff, which might cause resentment.

Without access to relevant data, it is difficult to assess the importance of variability in potential profitability; but intuitively, one might expect an innovation of this sort to be most profitable where other stages of the production process are modernised and in the construction of relatively sophisticated ships. Ship repairing firms have been excluded from the feasible set of firms which, in the event, is defined as the 62 shipbuilding firms quoted in the Geddes Report. [6]

(x) <u>Three innovations in weaving (EH, ADH, ASB)</u>.

(The only source of information for these three innovations, all in Lancashire weaving of cotton and allied textiles, is Metcalfe's own work.[1])

a) All three apply to the sizing process (a preparatory stage, prior to weaving, in which the warp threads are impregnated with size - this is necessary to prevent the threads from disintegrating in weaving.) The electrical hygrometer (EH) is a simple instrument which helps the operative measure the dryness of the warp; the accelerated drying hood (ADH) comprises a fan and motor covered by metal covers, its purpose is to increase the drying capacity of the 'tape frame'; the automatic size box (ASB) automatically controls the proportions of size paste and water applied to the warp. The main advantage of EH is that it increases the speed of the process and thus reduces factor inputs (almost exclusively labour) per unit of output, so too for ADH; ASB is also labour saving, in that certain sub-processes (e.g. pre-boiling the size) are totally done away with, but it also yields important savings in the amount of size needed and increases the weavability of the warps.

In 1950, EH and ADH cost £125 and £350 respectively to install; in 1962, ASB cost £2,750. The first two, then,are relatively minor innovations and ASB would imply a moderate investment for most firms. Metcalfe's estimates of typical paybacks are based on a boom period for the first two (1950) but a year of rather low activity for ASB. Somewhat arbitrarily, I have adjusted these estimates to make them all rather more typical of the whole period; assuming ADH yielded typical annual savings of £350 and EH £125 (both at 1950 prices), their paybacks are 1.75 and 1.867 respectively (slightly less profitable than Metcalfe's estimates.) As ASB diffused in a period of generally low demand, 1962 might be considered a typical year and therefore Metcalfe's estimate of a payback of seven years has been accepted.

b) The common factor to all three is that they are the result of research and development by the 'Shirley' Textile Research Institute, but all three have been sold through engineering firms (under licence). EH dates back to 1935 and has been sold by a single firm ; ADH (1949) was produced by six different firms and in this case there was a £20 royalty payment per machine; ASB (1951) was also only marketed by one firm paying ten per cent royalty. Given the crucial role of the research institute, it is difficult to say anything about the level of competition; on paper, in two cases, there is a monopolist seller, subject, of course, to the agreement with the Institute.

c) Metcalfe claims that some firms are too small to have tape frames and since without tape frames, none of the innovations are necessary, the rate of return from adoption will be zero. However, I have not excluded small firms from the feasible set, as presumably the only reasons stopping them from using tape frames are economic (i.e. backward management or insufficient liquidity) rather than technical. At any event, one of Metcalfe's footnotes shows some very small firms having adopted. [2] The only other way in which scale would appear to play a role is in large firms' better ability to use the innovations more intensively.

There is little information about post-invention improvements in any of the innovations, save that ASB had early teething troubles which were ironed out and that a Mark III ASB (presumably improved) was introduced in later years.

Metcalfe attaches some importance to the effects of the trade cycle and, intuitively, one can envisage that with insufficient demand, the extra capacity created by adoption is of little use.

The feasible set of adopters has been taken as all weavers in the Lancashire area. It would have been preferable to consider the whole U.K. industry but Metcalfe does not present the necessary data; the difference is probably minor. It would seem that within this set, profitability of adoption varied with certain technical attributes e.g. EH is less suitable for use on nylon and acetate rayon warps due to their higher electric resistance.

Finally, certain problems in measuring diffusion were caused by the relatively high death rate of firms over the period. This is considered more fully in the following appendix.

(xi) Automatic Track Lines (ATL) in Car Manufacture.

a) The Automatic track line is a classic case of automation: 'it replaces some, or many, individual machines, makes it possible for several metal-working processes on one piece of equipment to be amalgamated and caters for the internal transport, from one process to the next, of the work in progress.' [1] It is basically a series of machines operating on a continuous supply of engine parts e.g. blocks. The main economic advantages are savings in labour and the removal of potential bottle-necks.

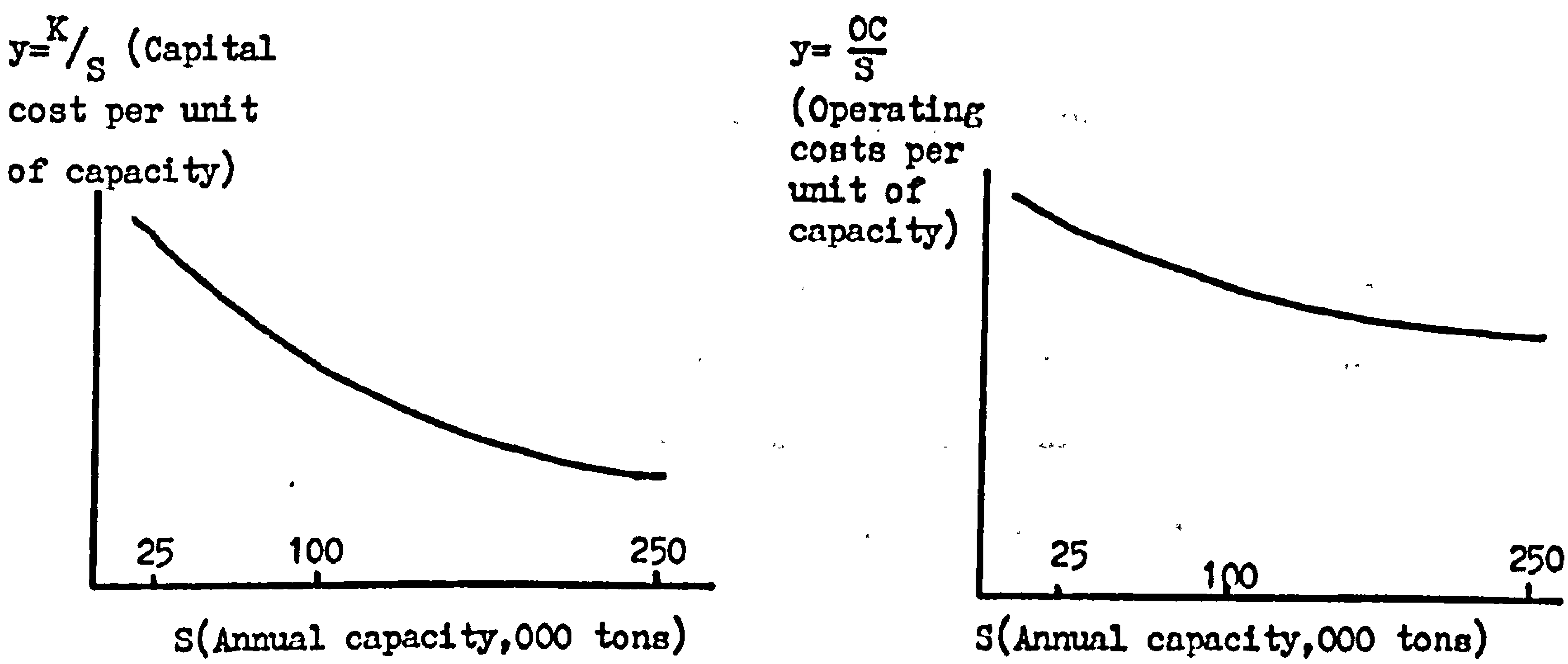
Even allowing for the large size of most British car manufacturers, the necessary installation costs are high, ranging from £650,000 for a line with an annual capacity of 25,000 units up to £9,000,000 for a capacity of 1,000,000 units. [2] Given the scale of early installations, an average cost might be £1,000,000 (at 1971 prices), although subsequent installations will have been more expensive. Four firms gave the National Institute estimates of the typical rate of return actually achieved, the average of which was twenty six per cent.

b) The first commercial introduction was in the U.S. after the war, followed quickly in various European countries, including the U.K. As opposed to virtually all the other innovations in my sample, this one originates from the car industry itself, although nowadays the installation of the lines is performed by engineering firms of varying sizes outside the industry.

c) Scale economies are quite pronounced: a chart produced in British Industry Week [3] shows this to be the most economical method of production once output exceeds 5,000 units; and data reproduced by Pratten shown in figure A.1.2. indicates scale economies in both initial costs and operating costs up to a relatively large scale (250,000 units.)

Fitting curves of the form $y=\alpha s^{B}$ to these two curves yields estimates of B of -.453 and -.206 respectively (the fit is very high, $R^2 = .95$, but there are only four observations for each !) suggesting that there are economies for scale of installation, particularly for the initial cost of the installation.

Figure A.1.2. Economies of scale for ATL



It is true that Pratten's data [4] suggests that scale economies cease above 250,000 units but it must be remembered that in 1966 (a date by which all firms had adopted), the top six producers produced only 1,600,000 cars and as some of these firms had more than one factory, this upper limit of 250,000 can rarely have been operative for any one plant.

The state of demand will have again played an important role in determining profitability as continuous operation is necessary to recover the capital outlay quickly and below full capacity working seriously impedes this objective.

As might be expected with such a major innovation, adoption implied fundamental internal re-organisation - 'major changes in the internal work - flow, including methods of tooling, works organisation and maintenance' [5] would be needed. Other disadvantages claimed [6] are high maintenance costs (especially in terms of skilled manpower), reduced flexibility of factory lay-out and high costs of breakdown.

Other lesser advantages are improvement in quality of product and reduction of rejects and scrap. The feasible set of firms has been taken as the six independent car makers existing before 1966 with capacities large enough to use ATL [7]. Given the standardised nature of product and process, there should have been little variance in profitability, abstracting from scale.

(xii) Computer Typesetting (CT) by Provincial Evening Papers.

a) This innovation is applicable to the 'composing room' in a printing works; it is, perhaps, self-explanatory - a computer is responsible for 'setting up' the page in terms of number of words per line, spaces between words and between lines [1] and, in this case, (newspapers) classified ads and display ads are stored, sorted and up-dated. These functions provide savings in composing room labour (e.g. reducing paste-up time) and improve the appearance of the newspapers. As composing staff comprised nearly twenty two per cent of the printing work force in 1967 [2], this can be seen as a relatively important innovation. Cost varies with the size of the computer used: in 1974 this ranged from £14,000 - £70,000 including soft ware, (i.e. mini-computers are used, that is, off line batch processing systems), although a few much larger computers have been installed which might cost up to £200,000. One of the manufacturers quoted a typical pay-back of two years i.e. a fairly profitable innovation. Again, a vintage approach would be inappropriate in this case, as no old technology embodied in existing equipment is replaced.

b) There is little point in tracing the development of the computer; however, with reference to this particular application, the first installations took place in the U.S.A. The first installation in the U.K. was in 1964 and there are three major producers in this country: Digital Equipment with the PDP8 and PDP11, Digico with the MICRO 16 and G.E.C. with the GEC 903; of course all of these computers have wide applicability, typesetting being only one of them. Two of these firms are American subsidiaries and each is large relative to its customers in the printing industry. Advertising literature is very sophisticated relative to that used for other innovations studied and after sales service is an integral and essential part of the product.

c) In communications received from the manufacturers, great emphasis was laid on the size of the potential adopter as a determinant of profitability – 'In all cases the shorter pay-back time is received by the larger company who can effectively use the computer for longer periods. The shorter the idle time of the computer, the faster it is cost effective.' [3] Over and above this, each system (say the PDP 8) encompasses a number of models which increase in scope and sophistication as size increases: to the extent that larger firms are more able to afford the larger models, they will benefit from greater returns which <u>may</u> reduce the pay back period.

Improvements in the technology over time have taken two forms: i) with successive generations of computer, learning by the manufacturer has led to increased facilities and ii) also presumably due to learning, price has fallen in money terms over time;[4] the typical price in 1973 was claimed to have been only twenty five per cent of the 1964 price.

Given the importance attached to avoidance of 'idle time' (see above) quite clearly the state of demand (especially in terms of quantityof advertising) will affect crucially the rate of return. There have certainly been problems in reconciling the introduction of computer-typesetting both with existing, often very outdated, techniques used on adjacent processes and with the work force in terms of re-training and opposition.

As computer-typesetting is not used at all by daily or London Evening newspapers – reputedly due to Union opposition, this study is confined to Provincial newspapers. As there are far fewer Provincial morning than evening papers, the latter has been taken as the feasible set; in fact most Provincial Dailies are printed physically at the same location as their 'sister' Evenings and use of CT on the latter implies its use on the former.

Local newspapers are excluded from the set as they are often far too small to use computer typesetting. Defining an industry of Provincial Newspapers presents difficulties, as, more often than not, they are not in competition with each other but with the National media. Nevertheless, this will only invalidate a small part of the model used (namely the lack of competitive pressures, resulting from other firms having adopted.)

(xiii) Tufted Carpet Machines (TC).

a) The tufted carpet machine is possibly the most fundamental innovation studied in this sample in that it revolutionises the basic process used and the type of end product produced by the carpet industry. Further, in most countries it has led to a basic change in the structure of the industry in that it has been the cause of substantial new entry. [1]

Basically, it transforms carpetmaking by replacing a weaving process with a giant sewing machine; pile yarn is inserted into a woven backing by a row of needles, the inserted tufts being held in place by the 'untwisting' effect of the yarn and by the addition of latex to the back of the carpet.

Although the cost of a machine is high (£35,000 in 1969 prices [2]) this comprises the largest and most important part of the carpet making process. Some idea of the labour savings possible can be gained from the 1968 Census of Production : operatives working on tufted carpet machines averaged 8,200 square yards of carpet per year, as opposed to only 2,100 square yards on traditional weaving machines. Using data from one of the manufacturer's advertising leaflets, an average sized T.C. machine, with an annual output of 165,000 square yards, might have a pay back, at 1968 prices, of 1.41 years.[3]

In many ways, this innovation fits well into a vintage approach.

b) Like most major innovations, TC machines have resulted from a series of independent patents leading on to development from a number of different sources. Reynolds suggests[4] that the largest single source of inventions in the historical development of the process were the textile producers in the U.S. As far as the U.K. is concerned, the innovation is supplied by four machine makers all situated in or around Blackburn, the most dominant of which are Singer (U.K.) Ltd., and Edgar Pickering Ltd. The latter, particularly, are heavily dependent on the carpet industry as a major market and have a number of subsidiaries overseas, claiming to be the world's largest exporters of Tufting machinery. Singer (U.K.) is a subsidiary of an American firm and is more diversified, having a separate Tufting Division.

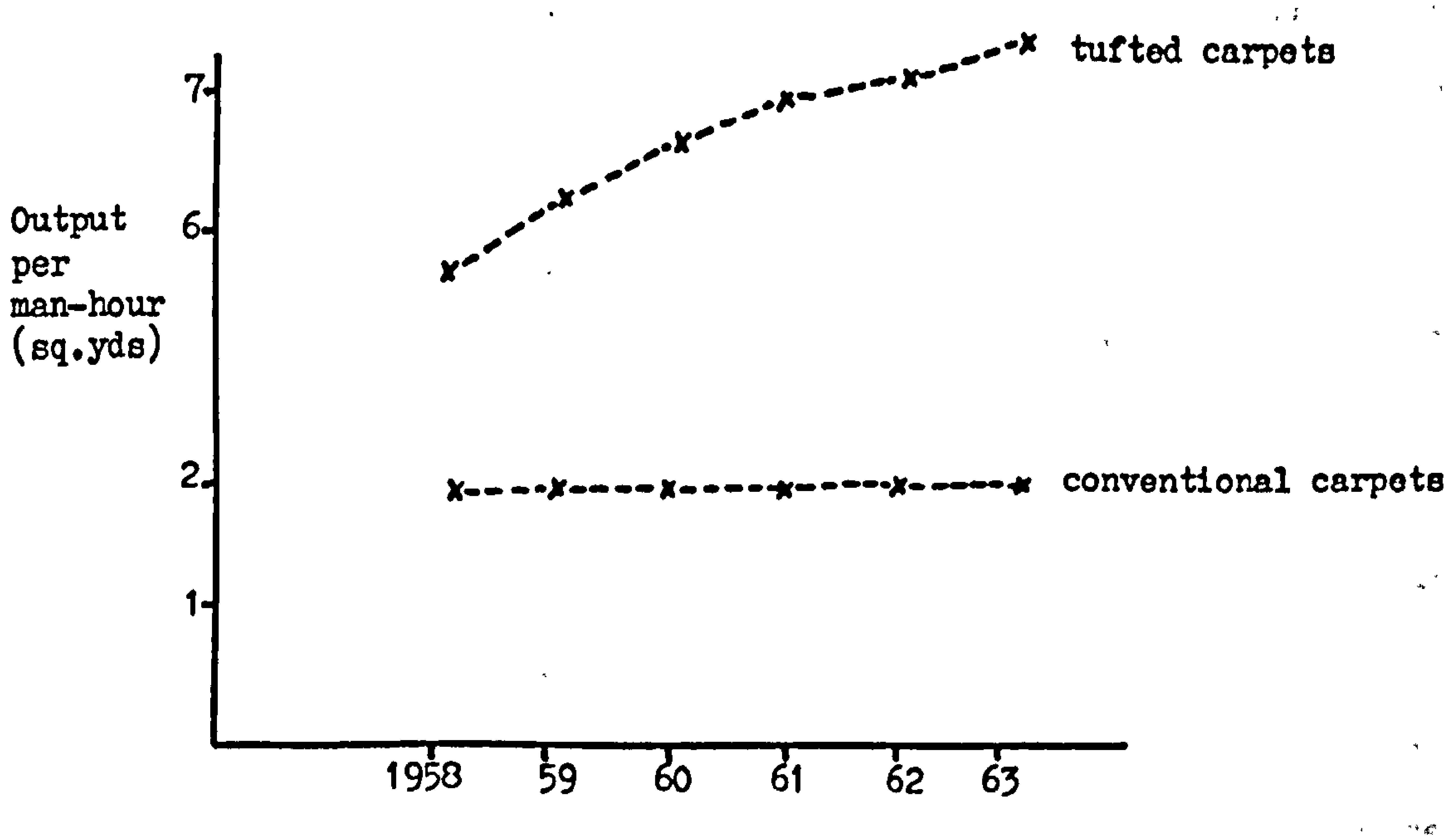
c) Evidence on economies of scale in adoption is somewhat enigmatic. T. Scott reports that the minimum optimum scale of adoption is at five or six machines with a combined width of twenty four yards. A combined width of sixteen yards (i.e. four machines) leads to average cost being 2.4% higher, and a combined width of eleven yards (three machines) has average cost five per cent higher. What the position is for even smaller scale of adoption is not clear. Whether most firms are, in practice, able to adopt at the twenty four yard level is doubtful, an investment of approximately £200,000 would be needed, However, Reynolds [5] talks of the effect of the tufting machine being to lower the size of the optimal plant because of its higher labour productivity. These two points are not in conflict and the second point does not eliminate the possibility of scale economies. The average size of carpet firms in 1963 was 270 employees and over three-quarters of all firms were below this size. It seems doubtful, therefore, whether many firms would be in a position to install sufficient TC machines to achieve optimum scale, and so, for the vast majority of firms in the industry, continuing economies of scale for adoption will apply. The numbers quoted above do not suggest, however, that they are very important.

Judging from manufacturers' advertising literature, there have been a series of vintages of machines, each an improvement on its predecessor.

Some data presented by Reynolds is suggestive on this score (Figure A.1.3): labour productivity between 1958 and 1963 in the U.S. increased much more quickly on TC machines than it did on conventional machines. This is due partially, no doubt, to learning by doing on the machines already installed, but almost certainly also to later vintages having higher labour productivity. As the slope of the curve showed no signs of decreasing over time, it is doubtful whether learning effects (which at the most might last for a couple of years) account for all of this increase. Disembodied general technical progress can probably also be ruled out as the conventional carpets show no similar trend in labour productivity. [6]

One factor which reputedly mitigated against speedier diffusion, at least in the early years, was the fact that only simple patterns could be produced

Figure A.1.3. The relative productivity advantage of TC in the U.S.



on TC machines. Indeed, certain firms with a highly differentiated product were able to refrain from adopting because they were shielded from the competitive pressures emanating from the lower cost tufted carpets. Recently, this factor has become less important as more and more sophisticated designs have become possible with later improvements in the technology.

As Reynolds suggests, [7] the cost of installation and profitability of TC machines will vary quite considerably across firms: 'There are so many variable factors that influence the production rate of machines ..... the type, size and gauge of the machine, the length of the stitches, the efficiency of the operator and especially the type and quality of yarn and backing fabrics.' There is, however, no reason why any existing carpet producer could not adopt. The feasible set has therefore been defined as all firms in the industry surviving through the period. New entrants have been excluded. [8]

(xiv) Tunnel Kilns (TK) in brickmaking.

a) There are three main stages in brickmaking : quarrying of clay, forming the clay into brick shapes and burning or baking the clay into the finished product in a kiln. The kiln represents the most expensive piece of equipment in any brickworks. The basic difference between the Tunnel kiln and the old technology is that in the latter, the raw bricks are stationary and the fire moves continuouslyaround the kiln, through a circular series of chambers, whilst in the new technology, the firing zone is always stationary, but the bricks move on a wagon at a very slow speed through the kiln, which is simply a long tunnel. [1] The main advantages of the Tunnel Kiln are: savings in labour, (primarily due to less physical handling of the bricks), more efficient use of fuel (in the old technology, fuel is wasted in building up the heat in each of the chambers separately), easier working conditions (in the old technology, workers have to physically enter the still hot chambers to remove the bricks after the fire has died down) and reduced chance of the kiln walls cracking.

An average sized [2] Tunnel Kiln costs about £260,000 in 1971 prices and the average payback given by twelve adopters in a N.I.E.S.R. survey was 5.6 years. Clearly, this is a very expensive and lumpy innovation (many brickmakers use only one kiln) which fits into Salter's vintage approach quite well.

b) Although the first patent for a Tunnel kiln was granted to a Dane as early as 1840 and tunnel ovens were being installed in the pottery industry (and a few early Tunnel Kilns also in the Brick Industry) before the Second World War, the technology, as it is known today, dates from the early fifties. It is rather difficult to attribute the effective source of this invention to any particular country, but the U.K. certainly played an important role.

A technical report by the Gas Council [3] lists eight separate Tunnel Kiln manufacturers, of whom at least two are foreign owned subsidiaries and three others are subsidiaries of leading process engineering firms in the U.K. For the latter, at least, the brick industry must represent a small part of their total market. Interestingly, at least two of the eight also build kilns of the old technology.

c) There is considerable evidence that TK exhibit marked economies of scale. Pratten suggests, on the basis of information from a leading manufacturer, [4] that 'substantial economies (of scale) continue up to an output of at least 25 million bricks a year for new kilns.' [5] Certainly, the figures in table A.1.3. confirm this. These scale economies would be of little importance if all firms were large enough to adopt at sufficiently high scale; however, this is not the case - the numbcr of firms with output at or exceeding 25 million bricks in the N.I.E.S.R. sample was only nineteen, and even these firms usually had a number of brickworks. Therefore the number of <u>plants</u> with output of above 25 million bricks p.a. is minimal.

Table A.1.3. Costs for New Tunnel Kilns

Annual Capacity (mn. bricks)	10	25
Total capital costs (£000)	184	296
Operating costs per 1,000 bricks (£):		
Fuel	1.78	1.50
Labour	1.27	0.86

(Reproduced from Pratten, Table 11.3, p.99)

Assuming that the scale curve can be approximated by the form:

$y = \alpha S^B$ where S = annual capacity, and y = capital cost per million bricks, B= -.32; if y = operating cost per million bricks, then B = -.13.

Moreover, multi-kiln firms have more chance of owning a kiln that is coming to the end of its physical or economic life, and may thus be able to derive maximum savings from adoption.

The sheer cost of the innovation might also mitigate against early adoption by small firms - they may have to wait longer to build up sufficient funds to finance the investment. Similarly, the increased technical expertise required by the relatively automated Tunnel Kiln may be beyond managers of smaller firms.

Crucially, Tunnel Kilns must be used continuously to derive maximum savings (they have relatively long start-up times etc.) and this naturally presents problems in periods of deficient demand. [6] The old technology is easier to slow down or even shut down for a short period, without incurring too much cost. Short run average cost curves, for both old and new technologies, would probably be U shaped with the minimum point lower for TK. However, one would expect the old technology to have a much shallower curve at least at lower capacities.

No information has been collected on the change in price of Tunnel Kilns over time, but there is some suggestive information about improved productivity. Davies and Smith show [7] that labour productivity in West German firms in 1969 is highest for firms who adopted a TK between 1964 and 1968, followed by those adopting in 1960 and 1964, and then pre-1960 adopters (and, incidentally, lowest of all for non-adopters.) This is not conclusive, as overall labour productivity depends upon many other factors, e.g. efficiency in employment of non-operatives, age of any other kiln used by the firms concerned etc., but it is at least consistent with the hypothesis of improvements from vintage to vintage.

At this point, one should mention the existence of the so-called 'sailing ship effect' i.e. since the advent of TK there have been improvements in the old technology, presumably in an attempt to combat the deterioration in relative costs for non-adopters. [8] However, for some reason, this seems to have been limited to Italy and there is little evidence of it in the U.K.

Finally, the nature of the product may be slightly different when using Tunnel Kilns: because the application of heat is more even, quality tends to be more uniform leading to fewer rejects. On the other hand, for just this reason, it is not possible to burn different types of brick simultaneously, which is sometimes possible under the old technology. This will clearly be a severe drawback for those firms producing relatively small quantities of a number of different typed bricks.

Definition of the feasible set of adopters is particularly difficult. Initially, at least, all clay brick firms are considered; but for technical reasons, many of these are ruled out. Most importantly, many firms use clay that contains a lot of carbon; if the carbon content exceeds a certain amount, then the old technology has lower costs, mainly because flexibility is needed in application of heat (because the carbon itself acts as a fuel). This flexibility is much less attainable in the Tunnel Kiln. This effectively rules out the fletton [9] sector of the industry and a number of other producers.[10]

However, carbon content varies throughout the country and, even amongst firms with a lesser amount in their clay (the critical level has been taken as 3%), although TK may still be possible, its potential rate of return will be reduced. Therefore, variable carbon content certainly produces variability in potential rates of returns, as does variability in product mixes.

(xv) The Basic Oxygen Process (BOP) in Steelmaking.[1]

a) The steelmaking process is only one part of the overall technology of the Iron and Steel Industry. To obtain steel from 'pig iron' the latter must be refined, namely, unwanted chemical elements (carbon, sulphur, silicon, manganese and phosphorous) must be removed by oxidation. (The pig iron itself, is produced from ore in blast furnaces and once steel has been produced, there are still a number of what might be called finishing processes before the end product is complete.) Traditionally, steel was refined using the acid and basic Bessemer process, in which air is blown from underneath into the hot metal bath in a converter; or in open hearth furnaces, in which burning fuel gas is passed over the top of the pig iron and scrap metal. [2]

The basis of the oxygen steel process is the use of pure oxygen (rather than air) which is blown through a water-cooled lance from above on the surface of the hot metal bath (again, a certain, more limited, amount of scrap metal may be used in addition to the pig iron.) There are a number of variants of this process: L.D-AC, VLN, Kaldo and Rotor, all of which are treated as one basic innovation in the following.

The main advantages of the new technology are shorter tap-to-tap time (i.e. quicker output from the same amount of metal input), improved quality of steel (particularly as compared with the Bessemer processes), much lower fuel costs and cheaper capital costs (compared to new O.H. or Bessemer furnaces.)

On the other hand, there is a relatively low limit to the amount of scrap that can be used with the pig iron, (whilst for O.H. furnaces fifty per cent of the charge may be made up by scrap, for Basic Oxygen, it must not exceed thirty per cent) and consequently, the furnace must be physically close to a blast furnace (providing its major material, pig iron). Both of these factors may have limited the diffusion of the new technology.

The investment needed to install a basic oxygen unit is high even relative to the average size of firm in the industry - a 'typical' cost in 1969 prices might be six million pounds. [3] Typical profitability of installation is much more difficult to measure, but the four adopters giving N.I.E.S.R. estimates achieved a mean payback of six years exactly.

b) The technology proper probably originated in Linz, Austria just after the war, After overcoming a number of teething troubles with pilot plants, in late 1952 a very modest first commercial installation was made, followed by a similar plant in Donawitz (also Austria). These represented the first installations of the so-called L-D variant of the process, other variants originated slightly later in different countries : LD-AC in Belgium, Rotor in Germany, Kaldo in Sweden and the Ajax process (really the conversion of an open hearth furnace) in the U.K.

As far as the developments in this country are concerned, no one source was responsible, to a certain extent the Research Association, Process Engineering firms and the Steel firms themselves collaborated.

No installation is exactly the same as any other and so the planning stage always involves a certain amount of development work peculiar to that installation. There are four major process engineering firms who install basic oxygen process in this country. [4]

c) There are certainly economies in the capital cost and operating costs of basic oxygen as the scale of the installation increases. Before considering these, however, it is important to note that there has already been a fair amount of work, both by economists and engineers, on scale economies [5] for any plant or works in the steel industry, with the general conclusions that

$I = a_1 S^x$ .... (i) where I = initial investment cost of the equipment or plant

$OC = b_1 q + b_2 S^x$ ....(ii) S = plant size (i.e. throughput at full capacity)

OC = operating costs, $b_1 q$ = fully variable costs, such as fuel and any tonnage-based bonus element in labour costs, (q = actual output) and $b_2 S^x$ reflects fixed labour and other costs.

Where (i) and (ii) have been estimated for plants $2/3 < x < 3/4$ and
" " " " " complete works $.7 < x < .9$

One might expect these economies to obtain for the basic oxygen process (and, for that matter, the three other innovations in the steel industry considered subsequently). Leckie and Morris have estimated the cost curves specifically for basic oxygen in some detail, taking into account not just rated plant capacity but also vessel sizes (any given capacity can be achieved for BOP in a number of ways, essentially by varying the number of converters used in the installation and the size of each converter.) Their cost curves reflect the optimal choice at each capacity; for instance, for 40,000 tons per week, it is optimal to install two converters with capacity of 200 tons each, but for 20,000 tons, it is optimal to employ three of 60 tons each. Algebraically, they find the following relationship yields a good explanation of capital costs: $K = n135000M^{2/3} + 95Q + 2140Q_c^{2/3}$ where K = investment cost, n = the number of converters used, M = the nominal size of the vessel (tons), Q = output from whole plant (tons per week), Q_c = capacity from whole plant (tons per week).

The graphical counterpart of this equation (figure A.1.4.) using, for each output, the optimal mix of size and number of converters shows obvious scale economies which can be approximated by $K = \alpha S^{.8}$.

Analysing current operating costs (1967) for the innovation yielded the following approximation: $R_q = \frac{n(178000+2200M)}{Q} + 34$ with M, n and Q as before, R_q = operating costs in shillings per ton.

Graphically, this portrays a classic L shaped cost curve becoming completely horizontal at a capacity of two million tons per year (which is in excess of any BOP installation to date.) Therefore, over the relevant range, unit operating costs can also be approximated by $OC = \alpha S^B$ where B is estimated as −.3, (very similar to the figure implied above for scale economies in capital costs per unit of output of −.2).

Other factors working in favour of large firms are their presumed greater ability to raise the exceptionally high capital sums needed and the higher probability of having the most relevant processes and products for basic oxygen.

Development of the technology has continued over the years in a number of directions. The feasible product mix has been widened - initially BOP was most profitable in producing common grade carbon steels and for processing hot iron with a relatively low phosphorous content, but by the early sixties

Figure A.1.4. Economies of scale in capital costs for BOP.

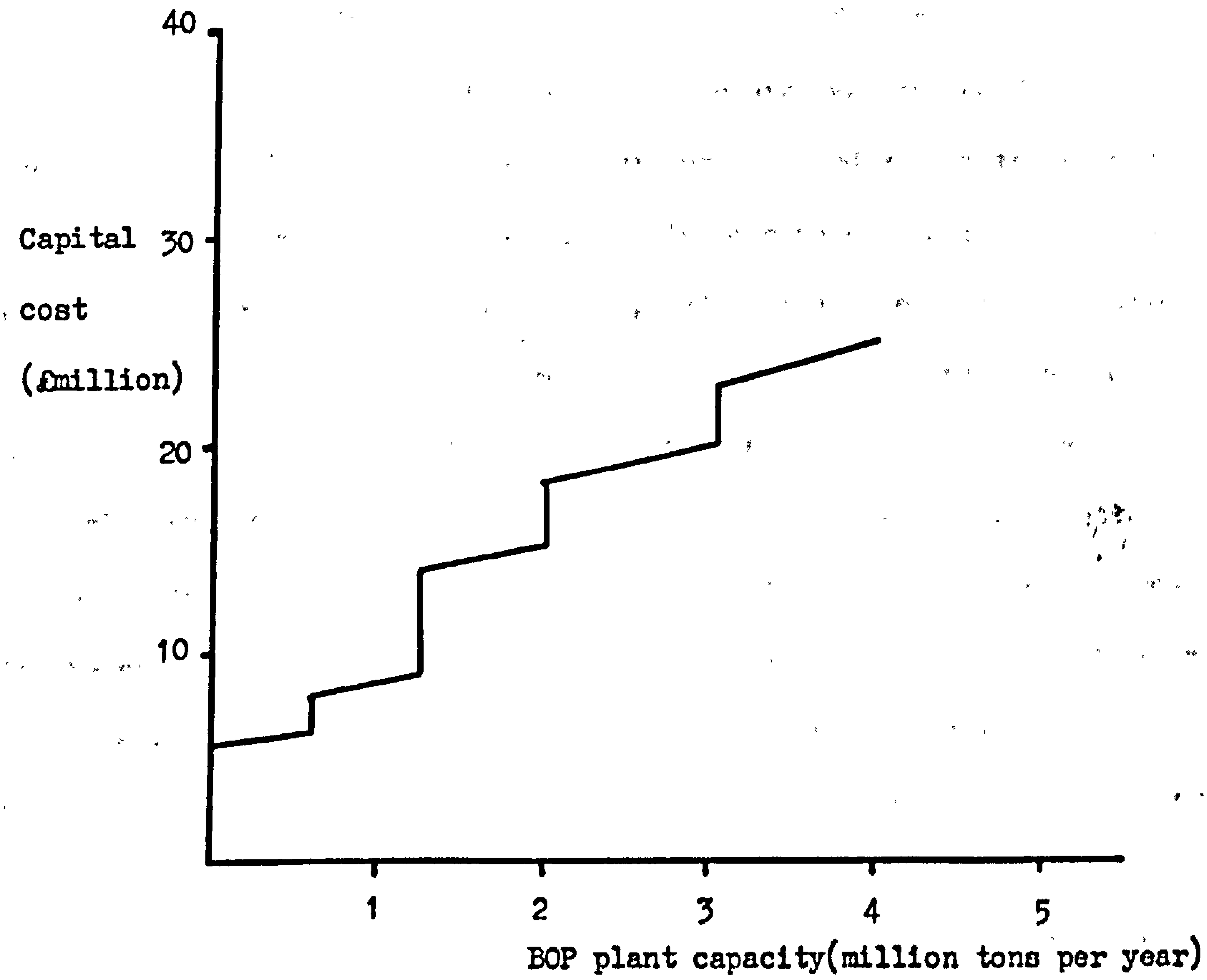
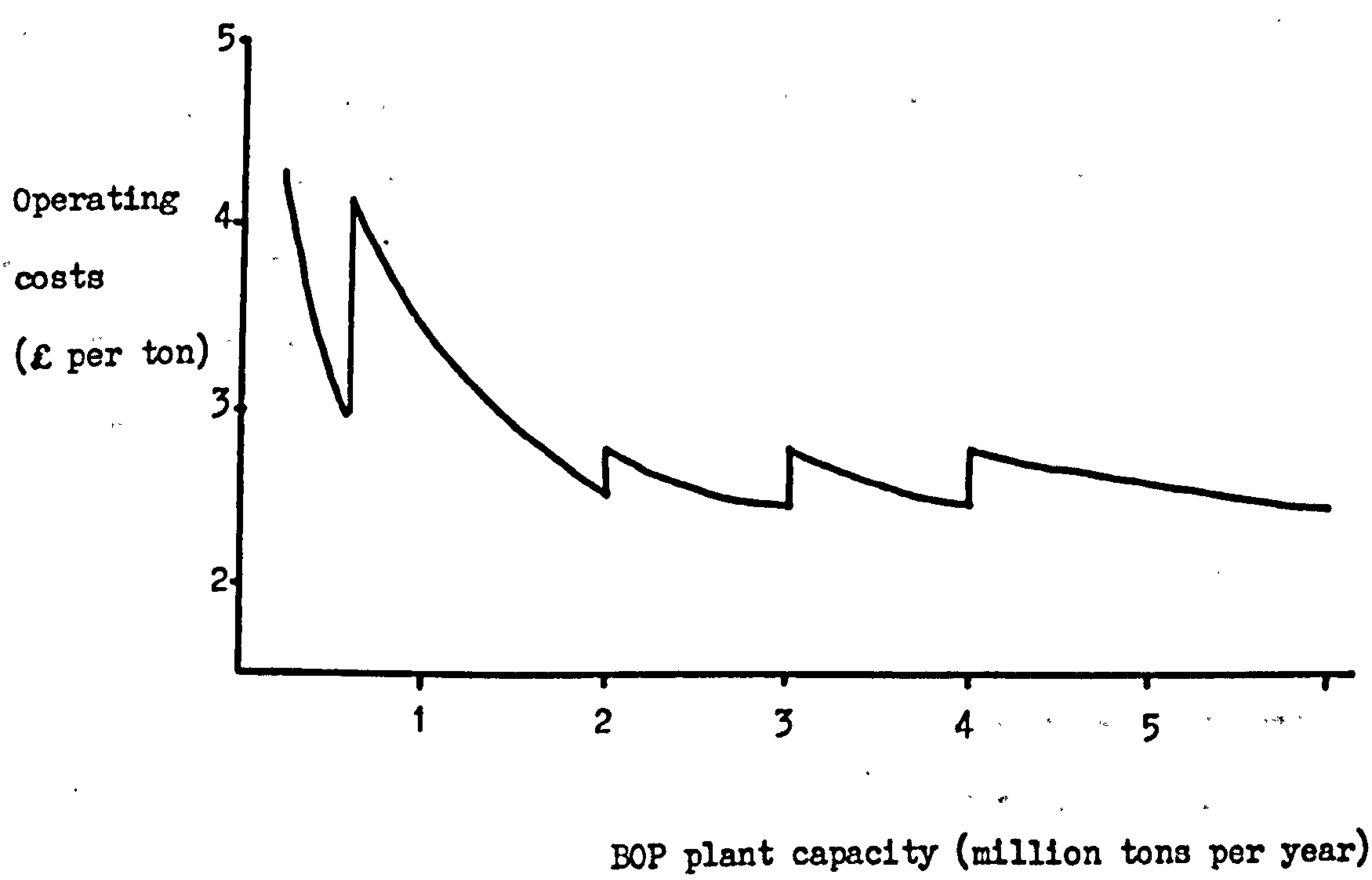


Figure A.1.5. Economies of scale in operating costs for BOP.



it was also possible to use BOP for high carbon and alloy steels and high phosphorous ores could be used with the addition of lime. The maximum amount of scrap which could be used in the charge was raised from 20% to 30% and it is possible that this may be further extended with more research.[6] At least in the early years, most installations were on a relatively small scale (the innovating country, Austria, had no real need for very large furnaces because of its limited market) but, as the innovation spread to other countries, scale of installations rose and at the same time, economies of scale were established. From the U.K. standpoint, most of these improvements were made before the innovation started to diffuse. However, to the extent that improvements are ongoing (especially in terms of learning by doing), one might suppose that the U.K. industry is faced with an innovation still subject to vintage to vintage improvements, if to a lesser extent than was true in the early years.

Interestingly, the old technology has also undergone technological development since the introduction of BOP; open hearth furnaces are now sometimes oxygen enriched (i.e. oxygen is mixed with the fuel to increase flame temperature and fuel firing rates.) It should be stressed, however, that these improvements have been made on existing OH furnaces - there have been no new OH installations since the advent of the Basic Oxygen Process.

Finally, a number of obvious points can be made:

i) because faster output rates are one of the major advantages of BOP, demand and capacity usage must be high, otherwise there will be little point in increased speed and savings in some semi-fixed costs per unit of output will be reduced.

ii) installation of BOP implies a reasonably long period of disruption in the works: the technology is radically different, necessitating fundamental re-organisation of the entire works.

iii) in some cases, the quality of the end product is improved (a number of firms cited this as a major advantage when answering the N.I.E.S.R. questionnaires.)

Two problems have to be faced in assessing the number of firms who were potential adopters of this innovation: i) how to handle nationalisation ii) what allowance should be made for the technical limitations of BOP. The period during which diffusion has been studied has been set at 1958 to 1968, which, given a typical time lag of three years between the start of the decision making process and the actual first use of BOP, represents a period (1955 to 1965) ending at the time of nationalisation. Fortunately, the second problem is also easily resolved: no firm without a blast furnace has or could adopt BOP, given the necessity for the use of hot charge in the latter; the feasible set can be simply defined, therefore, as all firms using blast furnaces.

Even within this feasible set, however, potential profitability of adoption will vary depending on existing steelmaking techniques, nature of end product and raw materials used.

(xvi) Continuous Casting (CC) in Steelmaking.

a) Following the refining of steel, the next process in the Iron and Steel industry is the casting of the molten steel traditionally into ingots, which are then rolled into semi-finished products, such as billets, blooms or slabs. (The following stage is the conversion of these into finished products by further rolling and other related processes.) This traditional method of casting is essentially intermittent: after each ingot group is cast, casting is stopped in order to 'trim off' large quantities at the top and the bottom of each ingot; after the metal is cooled sufficiently, the ingots are removed from the moulds, reheated and then transferred to 'blooming mills', where they are rolled into the semi-finished product. Continuous casting [1] turns this process into a continuous one: the liquid metal is poured into the mould at the same rate as the solid metal emerges at the other end, to be passed through straightening rollers and cut into pieces of the required length. The main advantages are the total displacement of the soaking pit and the blooming mill and improved yield (mainly because of the displacement of 'trimming off'). An average sized CC machine (producing 300,000 tons of six inch square blooms per annum) in 1964 cost £4,640,000 [2] and although it has proved difficult to derive reliable estimates of the typical payback, the same source optimistically calculated annual savings for the above machine at about £1,000,000. An average estimate from <u>actual</u> installations (from N.I.E.S.R. questionnaires), however, is of a pay back period of 7.73 years.

b) Continuous casting of non-ferrous metals dates back to the late 1930's in Germany and the U.S. but the special problems [3] of applying the technique (which was 'thought of' early last century) to steel were not overcome in commercial application until 1952 in Austria. In the U.K., experimental work was started up by B.I.S.R.A. (the research association) in 1947 [4] but the first commercial installation occurred in 1960. There has been a need for experimental lines for most firms, who have ultimately adopted, in order that they may determine how applicable CC is to their own steelmaking plant and products. Nine engineering companies are known to install CC lines, although presumably most of their work, historically, has been for the non-ferrous metal industry. As with BOP, it would not be totally accurate to portray the position as a set of engineering firms selling an innovation to the steel firms. The technology is really developed, or at least widened, with each new installation and, as such, each adopter is helping to improve the technique. On the other hand, the concentration of knowledge about the innovation is very much in the hands of the engineering firms, [5] if only because of their wider experience through past installations.

c) Ray [6] claims that CC has reduced the 'minimum economic size' of a steel works because it removes the need for blooming mills which are always, by necessity, large in terms of necessary capacity. As CC can be afforded by smaller firms, they are now able to use the best practice technique, which was not possible before. This is not to say that there are not economies of scale in the usage of CC, but rather that the old technique to be replaced is not common to all firms. As such, it is possible that smaller firms will derive bigger gains from adoption. [7] There appears to be no published data on the scale economies of CC, but there is no reason to suppose that it differs radically from other steel plant. Thus, one might expect to find that 'the higher the capacity of a unit, the lower are the capital charges, because capacity tends to be a cubic function of linear dimensions, whereas cost tends to be a quadratic function.' [8]

One factor possibly leading to higher profitability of installation for larger firms is that they are more likely to use the most appropriate adjacent equipment and/or produce the most appropriate product (simply because of their size, one would expect them to use a wider range of processes and produce a wider range of products.) Gebhardt suggests, for instance, that CC is more profitably combined with the Basic Oxygen Process than with Open Hearth Furnaces [9] (it is most efficient when the charges of liquid steel are available, as they are with BOP, at regular intervals within rather narrow limits.) Further, until recently, only billets or small blooms could be produced by Continuous Casting.

Improvements in the technology have taken two main forms: i) as just mentioned, it is only recently that some types of semi-finished product could be produced by CC i.e. the scope of the innovation has widened, ii) the economies of large scale introduction have been increased over time and this is reflected in a constant rise in average capacity of installations.[10] More generally, teething troubles have been ironed out only gradually, and then only due to continued experimentation by the research association and certain firms. Ray [11] maintains that the process 'is still in a phase of continuing technological development.'

Given that it embodies such a major change from a batch to continuous process, he adds that 'apart from these (high investment) costs, the primary deterrent for most firms is the need for reorganisation if proper co-ordination of existing melting, casting and rolling capacities is to be achieved.' By the same token, being a continuous process, it is more necessary than usual for the innovation to be operated at something approaching full capacity for any real gains to ensue.

Although much of the earlier development work was intended for application to special steels, it has not resulted in any actual installations producing these steels in the U.K. (and very few in the world as a whole.) Gebhardt suggests [12] that the reason for this apparent non-applicability is the heterogeneous nature of most special steelmakers' output. The continual stopping of the processes needed to produce a number of different steels would substantially reduce its actual operation time and, with it, any real cost savings, from adoption. Following Gebhardt's suggestion, therefore, all special steel makers have been excluded from the feasible set of potential adopters.

But as the U.N.E.C.E. report on continuous casting confirms [13], cost savings, even within the feasible set will vary considerably: '(they) can be affected by a great many factors relating peculiarly to that particular works - the location of the works, its layout and the general complex and details of production processes......., its general production programme and the size of its mrket, the supplies of operating materials (etc.)'

(xvii) Vacuum Melting (VM) in Iron and Steel.

a) Vacuum Melting is a process for producing or refining very high grade steel. Traditionally, the main source of high quality alloy engineering steels has been the electric arc furnace.[1] (In fact, in recent years, this has also been used in place of open hearth furnaces for the production of bulk steel.) However, for certain types of specialist steel, the arc furnace is not capable of producing the refinement really needed; in the past therefore, steelmakers have had to be very careful in selecting raw materials used and final product offered to their customers. Basically the problem is that such steels must be able to cope with extreme temperature and high stress, which they may not be able to do, if they are impure (i.e. include oxygen, hydrogen, sulphur etc.) Unfortunately, impurities are often too high in arc furnaces, partly because the materials are melted in air. There are three main variants of vacuum melting: vacuum induction melting, vacuum remelting and electroslag remelting. In basic vacuum induction melting, the charge (raw materials) are placed in a furnace which is, itself, installed inside a steel chamber that is sealed and evacuated just before melting starts. Of course, none of the gases in the air can get into the chamber and those unwanted gases present in the mixture are liberated on heating and drawn out by continuously-running vacuum pumps. There may be a second chamber in which the molten alloy can be cast into ingots or castings without breaking the vacuum. This method is typically used for small quantities of steel. The consumable electrode vacuum remelting process consists of a first stage, in which the steel is refined in contact with air and cast into a large round bar, which is then used as an electrode in a vacuum furnace. An 'arc is struck' between it and a small pool of molten metal in a water cooled mould. As melting proceeds, the original electrode melts away and the metal builds up below the pool in the mould (unwanted gas having been drawn off.) Electro slag remelting is rather similar, except that the heat is supplied by the passage of current in the medium of an electrically conducting slag. Typical installation costs are given in 1969 prices as £150,000 - £200,000 [2] although there may be substantial variation around and within this band. It is almost impossible to calculate a typical payback period, as one of the major incentives for adoption is the improvement in product. To a certain extent, the same quality product can be produced using the old technology but, as has already been stated, this requires careful selection of raw materials and high rejection rates of the finished product. To the extent that this is avoided using VM, it is considered that a rough estimate of the cost savings [3] allowed would produce a payback of about five years.[4]

b) The need for a high quality probably acted as a stimulus for the development of a series of innovations [5]; it was soon established that vacuum degassing (see following section) was not generally capable of producing sufficient refinement. Initially, vacuum induction melting was used (in fact a German producer was using this from 1923 onwards) but unfortunately this could produce only very small quantities; the next stage - the introduction of consumable electrode vacuum re-melting was almost by accident. It was originally introduced (both in the U.K. and U.S.) for ingots in titanium, but excess capacity led to trial remelts of steel which proved successful. Plants for remelting steel began to be established from 1958 onwards. Finally, electroslag remelting (based on a process first invented in the U.S. in 1939) appeared in this country in 1962, although it was already being used in the U.S.S.R. some time before. In the meantime, the original process (induction melting) has been improved and is a viable alternative for some uses.

Excluding two steel firms who installed their own VM units (one of them in co-operation with the research association), there are eight firms who have been responsible for installations in this country, all of them subsidiaries of major process engineering firms. World wide, including the U.K., there are over thirty manufacturers.

c) There is no specific published data on the economies of scale for this innovation. For the remelting variants, however, unit costs apparently decline with the diameter of the ingot to be remelted, and as this in turn is related to the size of the original melting part of the process, it seems likely that there are large scale economies in operating costs. More generally, one would expect the economies in capital costs found for all sorts of chambers or containers in the steel industry to apply. (See section on basic oxygen.)

There have been continual improvements in the technology, most particularly in the form of the improved product quality and Barraclough suggests that electroslag units will continue to increase in size, based on research findings from the U.S.S.R.

The feasible set has been defined as all firms producing 'permanent magnet, tool and high speed steels' as defined by B.I.S.P.A. This grouping certainly includes all firms for which VM is applicable; it may just include some for which it is not. Virtually, all production is from the private sector of the industry; fortunately those firms that have since been incorporated into BSC within this group had already adopted before nationalisation and hence are considered as separate firms.

Within this tightly defined product grouping, there are still differences between products which affect the profitability of adoption for different firms.

(xviii) Vacuum Degassing (VD) in Steelmaking.

a) To simplify crudely, vacuum degassing provides the same function for bulk steelmaking that vacuum melting does for those special steels considered in the last section, i.e. the liberation or removal of unwanted elements (usually gases) from the finished ingots. One important difference is that Vacuum Degassing units are often used in conjunction with the conventional methods of melting and refining (i.e. electric furnaces, open hearth and basic oxygen converters.) The advantages are mainly concerned, again, with the quality of the end product: better hydrogen removal avoids lengthy and costly heat treatment and reduces the tendency of the steel to hair-line cracking, the steel also has better properties (hardenability, weldability, fatigue etc.) Productivity is improved due to a reduction in subsequent processing (elimination of heat treatment and eradication of the need for slow cooling) and an increased ability to use cheaper bulk low carbon steels to make sophisticated end products.

The simplest form of VD unit - ladle degassing - consists of a ladle containing a hundred tons or more of molten steel being lowered into a large steel vessel, a cover is then clamped mechanically on top; very large capacity pumps evacuate the degassing chamber and hold the vacuum for up to twenty five minutes, causing unwanted gases to be thus liberated and drawn out. Variations on this theme involve the injection of inert gases and electromagnetic stirring to improve the contact between the metal and the vacuum. Other processes draw the molten metal itself up into an evacuated chamber or degass it as it passes from one ladle to another.[1]

The capital cost of installation varies greatly depending on site conditions and the intended end-use, but using data provided by plant manufacturers, for a 68 ton ladle capacity (the mean size of installation to date), the cost in 1969 prices would be about £164,000.[2] Cost savings can be substantial: one reference gives them as high as £25 - £50 per ton,[3] largely as a result of the reduced cost of heat treatment. A systematic appraisal of the costs of various VD units was made in an internal report of the BSC, the upshot of which is that such savings are probably rarely achieved; bearing in mind the typical capacity usage so far in the industry, an estimate of average payback would be 6.94 years.

b) A special report by the Iron and Steel Board chides British steel-makers and plant manufacturers for not undertaking very much early research on Vacuum Degassing and suggests that the main reason for this was their belief that the equipment necessary for experimentation was too expensive. [4] This did not apparently deter the American, German and Russian industries, which were responsible for nearly all of the initial research and development. [5] Consequently, until the early sixties, customers who required forgings made from vacuum degassed steel had to get it from abroad. The gloomy picture painted in that report seems a little unfair - certainly the first installations were in Russia in 1953 and Germany in 1954, but there were also two early installations in the U.K. operative by the end of 1956. On the whole, however, both the U.S. and Germany (and to a lesser extent Japan) were quicker to introduce and improve the new technology.

To date in the U.K., there have been eight firms responsible for installing VD units, two of them accounting for two thirds of the total (one a large German based international firm and the other a subsidiary of a large British process engineering firm.) However, both English Steel and G.K.N. installed their own units.

c) There are certainly economies in the scale of adoption of VD units. Diagrams based on data provided by plant manufacturers are presented by Holden, [6] relating cost of installation to ladle capacity and operating costs to output per year for two of the more advanced forms of VD (he claims that the shapes of the curves are very similar for other variants.) The data is reproduced here in table A.1.4., (a range is given in the original article but here the mid-point of the range only is reproduced.)

Table A.1.4. Economies of scale for Vacuum Degassing

A)	Output (000 tons per year)	50	100	200	300
	Direct Costs (£ per ton)	1.025	.725	.475	.4125
B)	Ladle Capacity (tons)	100	200	300	400
	Capital Cost (£000)	237	325	412	500

Fitting a curve of the form $y = \alpha S^B$ (where y = direct costs per ton and S = output) yields an acceptable fit with an estimate of B of -.5222. Capital costs, on the other hand, seem to be related to ladle capacity by a relationship more of the form $Y = \alpha + BS$ (where Y = aggregate, not per unit capital costs, and S = ladle capacity), with $\alpha = 150$ and $B = .875$. Or re-expressing in per unit terms : $y = 150/S + .875$, which does, of course, yield an inverse relationship between y and S, again with a gradient which declines (absolutely) as S increases (i.e. $d^2Y/dS^2 > 0$). However, the link between ladle capacity and annual output is not something that can be algebraically expressed, it depends on too many other factors peculiar to each installation. Suffice it to say, therefore, that there are also increasing returns to capital cost.

There are, in addition, a number of other factors mitigating towards earlier adoption by larger firms : possibly superior capability of financing the initial investment, higher probability of having an end product which would benefit most from VD, and perhaps higher probability of using BOP rather than OH furnaces [7] - the former being more compatible with VD.

There have been substantial improvements in the technology of VD over its first twenty years: the early plants used mechanical pumps but these were soon superceded by steam jet ejectors; generally, many operations that were initially manual have been mechanised; the scope of VD has been increased and, with it, the benefits of scale have become more pronounced.

This is largely due to the change in emphasis away from hydrogen removal towards deoxidization - consequently the average scale of installation has increased over the years.

Although in concept VD is really only 'an extension of accepted steelworks practice as far as metal handling, teeming and refractories etc. are concerned,' an internal report of BSC [8] suggests that few British installations have given completely satisfactory performance and cites as reasons, choice of degassing system and lack of technical competence, amongst others. This would suggest that VD is very much a fundamental innovation leading to a certain amount of disruption in the steel works.

The feasible set of firms is rather difficult to define. In the event, all firms using Open-Hearth, Basic Oxygen and Electric Arc Furnaces were considered potential adopters. Clearly, within this set of firms, potential profitability has varied considerably, depending crucially on the nature of the end product : in certain areas it has proved to be very successful (notably in the manufacture of medium and large forging ingots and in the production of certain alloy steels especially for bearings); whilst in others, although still profitable, it has been less of a success (heavy plate, sheets and some automative and aircraft steels.) Moreover, the type of VD unit used will depend on the method of steel refining in the works, as well as the nature of the end product, and it is apparent that certain types of VD have proved to be more profitable than others ex-post.

Firms that formed the BSC have been treated as separate entities; the period for which data was collected extends two or three years past the formation of BSC, but as the lag between the adoption decision and the unit coming into use is of about this length, the assumption of separate decision making is probably valid.

Technical References and Footnotes

For those innovations for which data was collected by N.I.E.S.R. there are two main references:

1. 'The diffusion of new industrial processes. An international study' edited by L. Nabseth and G.F. Ray (Cambridge University Press 1974.) - abbreviated in the following as 'New Industrial Processes.'

2. G.F. Ray, 'The diffusion of new technology.' National Institute Economic Review No. 48 May 1969, pp. 40 - 83. Abbreviated in the following as 'N.I.E.R. No. 48'

More specifically, the following footnotes apply to the foregoing technical appendix:

Special presses.

1. A more comprehensive description appears in 'New Industrial Processes' p.58
2. ibid: p.62
3. ibid: table 4.6 p.76 and table 4.7 p.77
4. ibid: table 4.7 p.77 and 4.13 p.83
5. ibid: p.82.
6. ibid. p.62.
7. 'The development of Special presses.' Dr. Nissan (The Robert Gordon's Institute of Technology, Aberdeen, Department of Paper Technology 1969)
8. 'New Industrial Processes' Chart 4.6 p.92, in which he uses the data differently to examine the relationship between pay back and propensity to actually install S.P.
9. See appendix 5 or chapter 7 for the implications of a straight line on logarithmic probability paper.

Foils.

1. As communicated by a leading manufacturer of foils. More generally, the major sources of information for this innovation were two manufacturers and the article quoted below.

2. G. Barnard, 'Foils, their efficiency on a slow machine.' Paper Technology, April 1971.

3. P. Hart, M. Utton and G. Walshe, 'Mergers and Concentration in British Industry' Cambridge University Press 1973. pp.151-3.

<u>Synthetic Wires (forming fabrics)</u>.

1. As communicated by a manufacturer. More generally, this manufacturer was a major source of information, as were E. Cruden and C. Wild,"Synthetic forming wires: progress design and development." The Paper Maker, April 1971 and P. Hampson, 'Wet felt economics', The Paper Maker, March 1971.

<u>Wet Suction Boxes</u>.

1. As communicated by a manufacturer, who was a leading source of information.
2. According to an anonymous article in the Paper Maker, Jan. 1962 p.32.

<u>Process Control of paper machines by computer</u>.

1. 'Paper Making' - an information pamphlet produced by Elliot Automation (1966).
2. For example, P. Stoneman (1974) op.cit.
3. 'Paper Making' op.cit. More generally, a series of articles in the 'Paper Maker' and 'Paper Technology' have proved invaluable sources of information.
4. This is the conventional finding in past work as reported by P. Stoneman p. 91 (op.cit.) but also see Stoneman for some econometric evidence suggesting that this might not be the case in more recent years.
5. Elliot Automation op. cit.

<u>Gibberellic Acid</u>.

1. For a more detailed description see 'New Industrial Processes' p. 215-6.
2. P.C. Northam, 'Brewer's Guild Journal' June 1962 Table VII p.302.
3. N.I.E.R. No. 48, p.80.
4. Communication from I.C.I.
5. N.I.E.R. No. 48, P.80.
6. 'New Industrial Processes' p. 227.
7. ibid.

Numerically Controlled machine tools.

1. For excellent descriptions of NCMT and long lists of their expected advantages over conventional machine tools, see 'New Industrial Processes', p. 24 and E. Mansfield, J.R. Rapoport, J. Schnee, S. Wagner, M. Hamburger, 'Research and Innovation in the Modern Corporation' (Norton, New York 1972)p.190.
2. 'Technical and Economic Aspects of Numerical Control'. J.R. Crookall (Machine Tool Trades Association, June 1968) p.11.
3. A pilot study (unpublished) for the N.I.E.R. 1968 article.
4. Mansfield et al, op.cit., p. 196.
5. N.I.E.R. No. 48, p.53 and 'New Industrial Processes' p.27.
6. 'New Industrial Processes' p.52.
7. N.I.E.R. No. 48 p.58.
8. 'New Industrial Processes' p.52.
9. ibid p.24.
10. ibid p.53.
11. N.I.E.R. No. 48, p.53.
12. 'New Industrial Processes' p. 41.
13. ibid. p.46 chart 3.6.
14. ibid. p.48 table 3.11.
15. ibid. p.49 table 3.12.

Shuttleless Looms.

1. For a more comprehensive description, see 'New Industrial Processes' particularly pp. 251-255.
2. There are four main types of shuttleless loom: Sulzer, Air-Jet, Water-jet, and Rapier.
3. 'New Industrial Processes' p.255.
4. A. Ormerod: Yorkshire Economic Bulletin 1963, Vol.15 "The prospects of the British cotton industry", p.15.
5. In response to an N.I.E.S.R. questionnaire.
6. In the empirical work, I have taken the average of these two (i.e. 16%) - the former estimate is undoubtedly too high, as it is based on very favourable

assumptions about shift working, the latter are little more than off-the-cuff guesses.

7. For a full account see N.I.E.R. No. 48, pp. 60-61.

8. 'New Industrial Processes' p.263.

9. ibid. p. 276.

10. N.I.E.R. No. 48 p.61.

11. ibid. p. 61.

12. 'New Industrial Processes' p. 269.

New methods of steel plate cutting.

1. For a more complete description see N.I.E.R. No. 48, pp. 72-3.

2. A reduction of 40% on labour costs at this stage of production is quoted in a paper by W.R. Mellanby in the Journal of The North-EastCoast Institution of Engineers and Shipbuilders, Vol. 75, 1958/59, p.267-8.

3. ibid.

4. In connection with their interim study reported in N.I.E.R. no. 48.

5. N.I.E.R. No. 48, p.75.

6. Cmnd. No. 2937, "Shipbuilding Inquiry Committee", 1965-66, H.M.S.O.

Three Innovations in weaving.

1. Manchester School, June 1970, op.cit. and his (unpublished) M.Sc. dissertation (Manchester).

2. Manchester School, op.cit.

Automatic track lines.

1. N.I.E.R. No. 48, p.76 (see also for a more detailed description of the technology)

2. Pratten, op.cit. p.138.

3. British Industry Week, 2nd February 1968.

4. Based on estimates from a U.K. motor company.

5. N.I.E.R. No. 48, p.76.

6. ibid. p.77.

7. ibid. p.77.

Computer typesetting.

1. Technically, they are used for hyphenation, justification, and formatting of text ready for the next stage of the printing process - hot metal or photocomposition machines. The major sources of technical information have been the machine makers, but also: 'Composing Room Controller', Data Systems April 1973 and Manpower Research Unit, "Report on Printing and Publishing" (Department of Employment), London 1970.

2. Manpower Research Unit op.cit.

3. Digital Equipment (20.3.1974).

4. For a more detailed consideration, see Stoneman op.cit.

Tufted carpet machines.

1. See T.W.K. Scott's (unpublished) Ph. D. dissertation (Sussex) and W.A. Reynolds, "Innovation in the U.S. carpet industry 1947-63" for detailed study of this phenomenon in the U.K., German and U.S. industries. Both of these references were used extensively in the writing of this section.

2. Estimate from T.W.K. Scott, op.cit.

3. This was confirmed as being reasonable by T.W.K. Scott.

4. Reynolds op.cit. chapter 6.

5. ibid. p.104.

6. Although these data refer to the U.S. industry, the machine makers are international and so one might expect a similar pattern in the U.K.

7. Reynolds op.cit. p.104: a quote from a leading machine maker.

8. The main point of the exercise is seen to be an explanation of how a given industry reacts to a new innovation. Including new entrants would entail an attempt at explaining why firms are able to enter certain industries more easily than others. Admittedly an interesting problem, this would necessitate a probably quite complicated extension of the existing model.

Tunnel Kilns.

1. For a fuller description see 'New Industrial Processes,' N.I.E.R. No. 48 and N.I.E.R. no. 58 pp. 54-71, 'The clay brick industry and the tunnel kiln.' S.W. Davies. Strictly speaking, the old technology referred to here is only one

(albeit the most important) of a number of older kiln types used in this industry.

2. With an annual capacity of 20 million bricks. The source for this figure is a leading kiln maker.

3. 'Gas in the heavy clay and refractories industries.' The Gas Council (Aug. 1968).

4. Pratten op.cit. p.99.

5. None of the firms in the N.I.E.S.R. sample had a kiln with capacity exceeding this level.

6. Many non-adopting firms, when asked for their reasons for non-adoption, mentioned the 'inflexibility of the T.K.' and 'depressed demand.' New Industrial Processes p. 114.

7. 'New Industrial Processes' Chart 5.2. p.112. In their work, this chart was not used to show this point.

8. ibid. p.108.

9. Fletton clay is peculiar to a part of England and is characterised by an extremely high carbon content.

10. See following appendix.

Basic Oxygen Steelmaking.

1. For more extensive discussions of this technology see N.I.E.R. No 48, pp. 41-2 and New Industrial Processes pp. 146-153.

2. There is also a third, more recent, process - electric furnaces - in which only scrap metal is used as a charge and the end product of which tends to be special steels. As such, it is not really competitive with basic oxygen. Since the war, the Bessemer process has virtually disappeared in the U.K.

3. For the average historical size of installation.

4. A small part of the world wide industry. There are about twenty nine companies overseas producing BOP.

5. Summarised in A. Leckie and A. Morris, 'Costs in the Iron and Steel Industry'. Journal of the Iron and Steel Institute, May 1968 pp.442-452.

6. New Industrial Processes p.151.

Continuous Casting.

1. For a fuller technical description see 'New Industrial Processes' pp.232-4 and N.I.E.R. No. 48 p.46.

2. B. Brisby et al, "Process selection in the steel industry", Journal of the Iron and Steel Institute, Sept. 1964 p.721.

3. Steel solidifies at a much lower temperature and is thus more difficult to keep liquid.

4. For a more detailed account of early British research, see 'Research in the Iron andSteel Industry: Special Report', 1963, Iron and Steel Board.

5. Who have sometimes undertaken joint ventures with B.I.S.R.A.

6. N.I.E.R. No. 48 p.46.

7. They will still be at a cost disadvantage, but less so.

8. W.F. Cartwright, 'The growth of unit output and its effects on works planning and management', Journal of the Iron and Steel Institute June 1969 p.729. See also the previous section on BOP.

9. New Industrial Processes p.238.

10. ibid p.240.

11. N.I.E.R. No. 48. p.48.

12. New Industrial processes p.238.

13. U.N.E.C.E. 'Economic aspects of continuous casting of steel' New York 1968.

Vacuum Melting.

1. A very extensive discussion of vacuum melting is presented by K.C. Barraclough, 'The newer specialist steelmaking and steel refining processes.' Journal of the Iron and Steel Institute, June 1969, p.826 onwards.

2. ibid p.835. (these figures have been converted from U.S. dollars.)

3. Strictly speaking, the immediate costs of production are higher but these are more than outweighed by the reduced risk of rejection at a later stage of what constitutes a very expensive product (in terms of pounds sterling per pound weight.)

4. According to an industry report.

5. Most of this account in particular is based upon Barraclough op.cit.

Vacuum Degassing.

1. For a much fuller treatment of the technical aspects and a more comprehensive delineation of the various alternative processes (that is stream degassing, circulation and vacuum lift degassing) see J. Flux, "Vacuum degassing of Steel," Journal of the Iron and Steel Institute 1965, p.1205, H. Holden, 'Vacuum degassing in the Iron and Steel Industry', Journal of the Iron and Steel Institute, 1969. Much of this section is based on these two references.

2. Holden, op.cit., p.809, chart 7.

3. ibid. p.810.

4. 'Research in the Iron and Steel Industry' op.cit.

5. D. Burn, "The Structure of British industry; a symposium", Cambridge University Press (1958) p. 484.

6. Holden op.cit. p.809.

7. This is borne out by the findings relating size of firm to frequency of BOP ownership (see chapter 7).

8. Report of the Vacuum Degassing Working Party, BSC Operating and Technical Standards Department, July 1971.

Appendix 2. The sources and quality of the data used in measuring diffusion.

As explained already, data has been collected from a variety of published and unpublished (as yet) case studies produced by other economists and from scientific and trade journals. [1] The first aim of this appendix is to establish the size of the feasible set of adopters for each innovation and, second, where the data is for a sample of firms only, to assess the applicability of the sample to the total population concerned. In section 4, time series diffusion data is presented graphically for the 22 innovations.

1. Sources.

The innovations can be grouped broadly into four categories:

a) those for which data is taken from trade and scientific sources (Continuous casting, basic oxygen, vacuum degassing, vacuum melting, computer typesetting and process control of paper machines by computer.)

b) those for which data was collected by other independent researchers (Tufted carpets machines by T. Scott and the three sizing innovations in weaving by Metcalfe.)

c) those for which data was collected by N.I.E.S.R. as a part of the International study of 'The diffusion of new industrial process'[2] (Special presses, gibberellic acid, tunnel kilns, automatic track lines, shuttleless looms, automatic steel plate cutting, and numerically controlled machine tools in the three branches of mechanical engineering.)

d) those for which data was collected by N.I.E.S.R. for the same study but which were not, in the event, used in the above work. (Synthetic fabrics, foils and Wet suction boxes.)

2. Determination of feasible sets.

As has become apparent from the previous appendix, most innovations have certain idiosyncracies which make them less appropriate to certain sectors of

1.But in all cases, it has been necessary to collect extra, mainly technical, information from trade journals, machine makers, research associations etc.

2. op. cit.

the adopting industries; in some extreme cases, these technical attributes effectively rule out certain firms from ever adopting the innovations. Clearly, in measuring diffusion, one would want to exclude such firms from the feasible set. In past work, this has led to concepts such as 'the technological ceiling' [1] or 'the technically feasible maximum'[2]. Mansfield overcomes this problem by studying only the largest firms in each industry (because in certain cases, 'it seemed very unlikely that firms smaller than this would have been able to use them'.) [3] Here, firms are only excluded from the feasible set if there are strong technical reasons for doing so; size itself is not considered reason enough.

The following innovations and industries are all straightforward; all firms within the appropriate industry have been included as potential adopters.

Table A2.2.1. Definitions of feasible sets

Innovation	Industry, and where applicable, source of firms' names.
Special presses	All paper and board makers as defined by B.P.B.M.A. reference tables
Synthetic fabrics	" " " "
Foils	" " " "
Wet Suction Boxes	" " " "
Process control by computer	" " " "
Numerically controlled machine tools	All manufacturers of Printing presses (Trade directories, 'Kompass,' N.I.E.S.R.-)
" " " "	All manufacturers of Turning machines(Trade directories, 'Kompass', N.I.E.S.R.)
" " " "	All manufacturers of Turbines (Trade Directories, 'Kompass', N.I.E.S.R.)
Automatic steel plate cutting	Geddes Report on Shipbuilding industry(op.cit)
Gibberellic Acid	All firms carrying out malting (Brewers' almanack, Kompass, N.I.E.S.R.)
Shuttleless Looms	All weavers as defined by the Textile Council
Computer Typesetting	All provincial evening newspapers (Kompass, Kelley's, Evening Newspaper Advertising Bureau.)

1. 'The diffusion of new industrial processes' op.cit. p.298
2. S.W. Davies (1971) op.cit. p.64.
3. Mansfield (1968) op.cit. p.135.

The other ten innovations and industries each have some technical points worth noting.

Basic oxygen steelmaking: all firms refining steel and using blast furnaces. (source: Iron and Steel Board Annual Statistics)

Continuous Casting: all firms refining steel but excluding solely-special-steelmakers - the same set as above, plus four additions (source: Iron and Steel Board, Annual Statistics)

Vacuum degassing: all firms using open-hearth, basic oxygen and electric arc furnaces (source: Iron and Steel Board, Annual Statistics.) As explained in the previous appendix, firms that were nationalised are considered as separate decision making units, as the period studied in each case stops at the date when investment projects were those decided upon by B.S.C.

Vacuum melting: all firms producing 'permanent magnet, tool and high speed steels' (plus one or two other special steels) as defined by B.I.S.P.A. (nationalisation is largely inapplicable to these firms.)

Automatic track lines: the six independent car manufacturers existing before 1960 (the date by which all had adopted) with capacities large enough to use ATL (as defined by N.I.E.S.R.)

Tunnel Kilns: all firms producing clay bricks excluding those using clays which make the tunnel kiln inoperable due to high carbon content. The list of all clay brick firms was constructed from trade directories, 'Kompass' etc; one problem that arises, however, is that the diffusion data from N.I.E.S.R. is available for only eighty firms and it is only for these firms that we know the carbon content of the clay. Consequently, it is assumed that the incidence of high carbon-clay-content-clay firms outside the sample is the same as for those in each of five size classes considered. This reduces the number of potential adopters to one hundred and thirty seven.

Tufted carpets: all existing carpet producers are assumed to be potential adopters. A special problem in this case, however, is that over the past few years, a number of new firms have entered the industry, in every case using tufted carpet machines. As has already been stated, it is considered best, for reasons of comparability, to exclude the new entrants and study diffusion only within the original set of carpet producers. According to T. Scott, in 1963 at least, all new entrants were in the size range 25 - 199 employees, therefore the feasible set of adopters is taken as those reported in the 1963 Census tables minus sixteen new entrants in the 25 - 199 range.

Three innovations in Lancashire weaving: Metcalfe presents [1] a series for the number of firms in Lancashire weaving since before the war, from which it is obvious that there has been a very large decline in population over those years in the industry. This will clearly affect any measures of diffusion used: for instance, should adopters who subsequently die be included within the overall figure for diffusion? In the event, diffusion is studied within that set of firms who have survived over the entire period. Thus all firms that have died during the diffusion period are excluded from the set of potential adopters, and those adopters who subsequently died are also excluded from the total set of adopters.

It should be remembered that diffusion is studied here only within Lancashire weaving which includes the overwhelming majority of British weavers.

One final point which should be made is that the number of potential adopters is calculated only for the final year for which diffusion data is available and so, implicitly, the assumption is that firm populations were constant over the period studied. [2] This is, of course, a simplification made because in many cases it would have been physically impossible to collect data for each year on the number of firms in each industry. (Very few of these 'industries' conform to the C.S.O's minimum list leadings exactly.) Casual empiricism would suggest that it may not be too unrealistic an assumption, especially given the relatively short period considered in many cases. As has just been mentioned, in the two industries where a fixed population was clearly not the case, special allowance has been made.

1. Unpublished M.Sc. thesis, University of Manchester, 1968.
2. Or that firms leaving or entering the industry included the same proportion of adopters as did the original surviving population.

3. Adequacy of the samples.

Most of the data collected by N.I.E.S.R. relates only to samples of firms within the appropriate industries; in this study, as one of the major aims is to compare rates of diffusion between industries, it is essential to examine the representativeness of these samples.

Of these twelve innovations, automatic track lines may be accepted immediately as all potential adopters were sampled. For the other eleven, although apparently selection was made by stratified random sampling techniques, the inevitable non-response must have introduced elements of bias. In many cases this was substantially reduced by repeated approaches to non-respondents (often with the result that such firms at least gave an answer to the question of whether or not they had adopted the new innovation - this is sufficient for my purposes at least.) In what follows, the sample results have been extended to the populations by stratification after selection and generally assuming no non-response bias.

Special presses: the sample set here comprises two separate samples carried out by the Institute (one for the purpose of an interim report and one for the study proper.) Although there was an 18 month gap between the two, the former sample were also questioned on whether they had plans to install within the following two years, and so the two samples may be considered as being taken in the same years.

Size Class [1]	N_i (number of firms in industry)	n_i (number of firms in sample)	m_i (of whom, adopters)
0 - 9999	51	4	0
10 - 19,999	15	9	3
20 - 49,999	18	18	10
$\geqslant$ 50,000	14	14	11
	98	45	24

1. tons of paper produced (industry size distribution from B.P.B.M.A. reference tables.)

N.B. In this and all subsequent tables in this appendix, n_i refers to respondents - there is no available hard data on non-response.

Because the two samples have been combined, coverage on the top 32 firms is 100%, and the only problem is that of the very poor coverage of the very small firm category. (However, according to industry experts, it is unlikely that any of these firms have yet installed special presses.)

The estimated number of firms that had installed SP by 1970 is therefore set equal to $\sum N_i\ (m_i/n_i) = 26$, and time series for the population's diffusion is hence derived by multiplying the time series for diffusion within the sample by a factor of 26/24. [1]

Foils, synthetic fabrics, Wet suction boxes: The sample set for each of these innovations was the same 24 companies (and coincides with the second sample mentioned for Special presses.)

		Foils		Synthetic fabrics		Wet suction boxes	
Size class	N_i	n_i	m_i	n_i	m_i	n_i	m_i
0 - 4999	35	0	0	0	0	0	0
5000 - 19,999	31	7	5	7	3	7	1
20,000 - 49,999	18	10	8	10	9	10	3
≥ 50,000	14	7	7	7	7	7	3
	98	24	20	24	19	24	7

Apparently non-response was extremely small in these cases [2] but a very serious drawback is the absence of sampled firms in the 0 - 4999 tons range. Ignoring this for the moment, the estimate of total number of adopting firms in the population would be 50.5; 45.5 and 15.8. Fortunately, in two cases, there are independent estimates (by machine makers) of the likely number of firms using the innovations: for Foils, 'probably slightly over 60,' and for Synthetic fabrics 'about 50'. These estimates would be consistent with say 11.5 users and 5 users respectively in the small sized firm ranges. (Happily such figures tie in well with what one might predict by extrapolating back

1. The same procedure for translating diffusion data for the sample to the whole industry is applied in the following nine samples.

2. As a test, the predicted penetration of SP, using this sample can be compared with that predicted by the larger sample of 45 firms - there was no difference in any size range. One might conclude that this set of 24 firms is representative of the larger set at least.

the Quasi-Engel curves of Chapter 7.) Consequently, it is estimated that for these three innovations, there were, in 1970, 62 ; 50.5 and 15.8 adopters (i.e. it is assumed that no small firm had adopted WSB.)

Gibberellic acid: 'all large and medium sized (malt makers) and a (random) sample of small brewers were approached[1],' according to the Institute. It does seem that non-response was small and was due mainly to closures or small brewers not actually making their own malt but 'buying out' from maltsters. The respondents have therefore been treated as a stratified random sample which leads to the conclusion that the state of diffusion in the sample by 1967, of 21 out of 31 having adopted, was equivalent to population figures of 36.275 out of 56 having adopted.

Size Class[1]	N_i	n_i	m_i
1 - 12	14	7	4
13 - 24	14	6	3
25 - 34	9	5	3
35 - 69	10	5	4
$\geqslant$ 70	9	8	7
	56	31	21

1. Measured by number of employees engaged either directly or indirectly in malting

Tunnel Kilns: as already mentioned, the feasible set of firms excludes fletton producers and those firms which use clay with a high carbon content.

Although N.I.E.S.R.'s sampling frame included such firms, they are excluded in the table below, which includes not only data on firms from N.I.E.S.R.'s sample, but also some that I have subsequently collected.

Size[1] class	N_i	n_i	m_i
1 - 24	58	12	0
25 - 49	33	17	5
50 - 99	19	10	3
100 - 249	16	8	4
250 and over	11	9	9
	137	56	21

1. Total employees.

1. N.I.E.S.R., No.48 op.cit. p.79.But note that they use an output definition of size.

Strictly speaking this is not a random sample; there was certainly non-response from firms of varying size but most non-respondents were in the 1 - 24 range. It is the conventional wisdom that, as far as is known, no firm in this range has adopted the TK; this is exactly the conclusion one would arrive at by assuming no bias in non-respondence.

Therefore, whilst still accepting that these respondents represent a potentially biassed sample, it seems fair to assert that such bias is unlikely to be substantial. 21 adopters in the sample, therefore is taken to represent 34.4 adopters in the population.

<u>Shuttleless looms</u>: as for special presses, the full sample here is the sum of two overlapping samples carried out in 1968 and 1970. A rough check is possible on the representativeness of the former as the Textile Council Quarterly [1] reports 1200 S.L's having been installed by that year.

First N.I.E.S.R. Sample.

Size class [1]	n_i	m_i	average number of looms per adopter	N_i^2	Estimated No. of looms
1 - 24	0	0	-	96	-
25 - 99	5	0	-	95	-
100 - 499	17	4	22	68	352
500 and over	11	7	64	20	896
	33	11		279	1248

1. Measured by total number of employees.
2. Hypothesised size distribution (see appendix 5, table A5.2.2. for the parameters of this distribution). Actual distribution unknown.

<u>If</u> the sample was representative, both in terms of proportion of adopters in each range and number of looms per adopter, there would have been 1248 S.L. installed by 1968 which does, in fact, exactly tally with the Textile Council's rounded figure estimate. It would seem, then, that the first survey, although subject to non-response, may have been adequate.

The second sample was mainly an attempt to acquire more information about the larger firms in the industry (including some that had already been sampled) and only adds seven previously unsampled firms (including 4 adopters in 1968.)

1. 'Textile Council Quarterly Statistical Review' 1969, The Textile Statistics Bureau.

Adding these in, yields 15 adopters out of 40 firms by 1968 - assuming that the first sample was indeed representative, this compares with an estimate of 30 adopters in the total population. Again, the residual question mark must be placed against poor coverage of very small firms. <u>Automatic steel plate cutting</u>: this sample probably represents most problems in interpretation for the whole population due to substantial non-response. The Institute split the industry into two groups: those firms capable of producing ships of more than 5000 GRT and those not. In the first group, 25 of 27 were sampled and in the second, 12 of 35; response was as follows:

	Total Population[1]	sampled	Responded	'Acceptable' non-response[2]
1st group	27	25	11	4
2nd group	35	12	3	3
	62	37	14	7

1. In 1965 (Geddes report.)
2. i.e. had gone bankrupt or been amalgamated by the date of the sample.

Further to this, two of the large non-respondents and one of the small non-respondents replied simply that they did not use the innovation, leaving 8 and 5 refusals to co-operate respectively.

To derive estimates of total penetration in 1967, the following steps are made:

a) to assume bankruptcy rates were similar in the firms not sampled which leaves, by 1967, 23 and 26 firms in the two groups.

b) to add the three firms who admitted they did not use the innovation to the respondents.

c) assume that the non-respondents in both strata had adopted to the same proportionate extent as the respondents.

d) assume that non-sampled firms in both strata had also adopted to the same extent (as all sampled firms.)

	Respondents	Respondents having adopted	Non-respondents	NR having adopted	not sampled	NS having adopted
1st group	13	8	8	4.92	2	1.23
2nd group	4	0	5	0	17	0

However reasonable are steps (a), (b) and (d), (c) is obviously arguable; quite possibly one reason for non-response may have been that those firms had not adopted. Replacing assumption (c) with one that no non-respondent had adopted, reduces the estimate for total adoption within the industry from 14.15 to 8.76 (assumed non-sampled adopters goes down to .76 as well.) However, it is considered that such an assumption is probably too drastic and so, in the absence of better information, heroically, an average of the two estimates is taken i.e. 11.45 adopters out of a feasible set of 49.

(This, of course, assumes that no firm with small capacity, as defined above, had adopted by 1967; this would appear to be the conventional wisdom.[1])

In conclusion, however, some doubt must be expressed against this estimate which can only be roughly accurate.

Numerically controlled machine tools: (a) in the Turning Machine industry.

A first sampling stage in this case involved telephone enquiries with virtual 100% success on the basic question of whether a firm had adopted or not; the fact that 100% coverage was not achieved must be due to use of a different sampling frame from the one that I have constructed [2] or incorrect addresses etc. These miscellaneous factors might cause bias in the effective sample but it is unlikely to be serious [3] and is ignored here. In other words, it is assumed that the 36 firms are a stratified random sample, in which case, 18 adopters in the sample implies 23 adopters in the industry as a whole.

1. According to a machine maker and N.I.E.S.R.
2. For the purposes only of estimating the firm size distribution.
3. Apart from under-representation of smaller firms which is allowed for by post-selection stratification.

Size [1] class	N_i	n_i	m_i
1 - 49	10	6	0
50 - 199	12	5	1
200 - 299	9	6	2
300 - 749	11	10	6
750 and over	11	9	9
	53	36	18

1. employees.

(b) in the Printing press industry.

The same comments as above apply here. Therefore, 10 adopters in the sample implies 17.4 adopters in the industry as a whole.

Size [1] class	N_i	n_i	m_i
1 - 74	13	8	1
75 - 499	18	7	3
500 and over	12	9	6
	43	24	10

1. employees.

(c) in the Turbine industry.

Again, the same comments apply as in the above two cases. It must be acknowledged, however, that both industry and sample size are small, thus reducing the reliability of the estimate for the whole industry.

Size [1] class	N_i	n_i	m_i
1 - 499	6	2	1
500 - 1999	7	2	1
2000 and over	7	5	5
	20	9	7

1. employees.

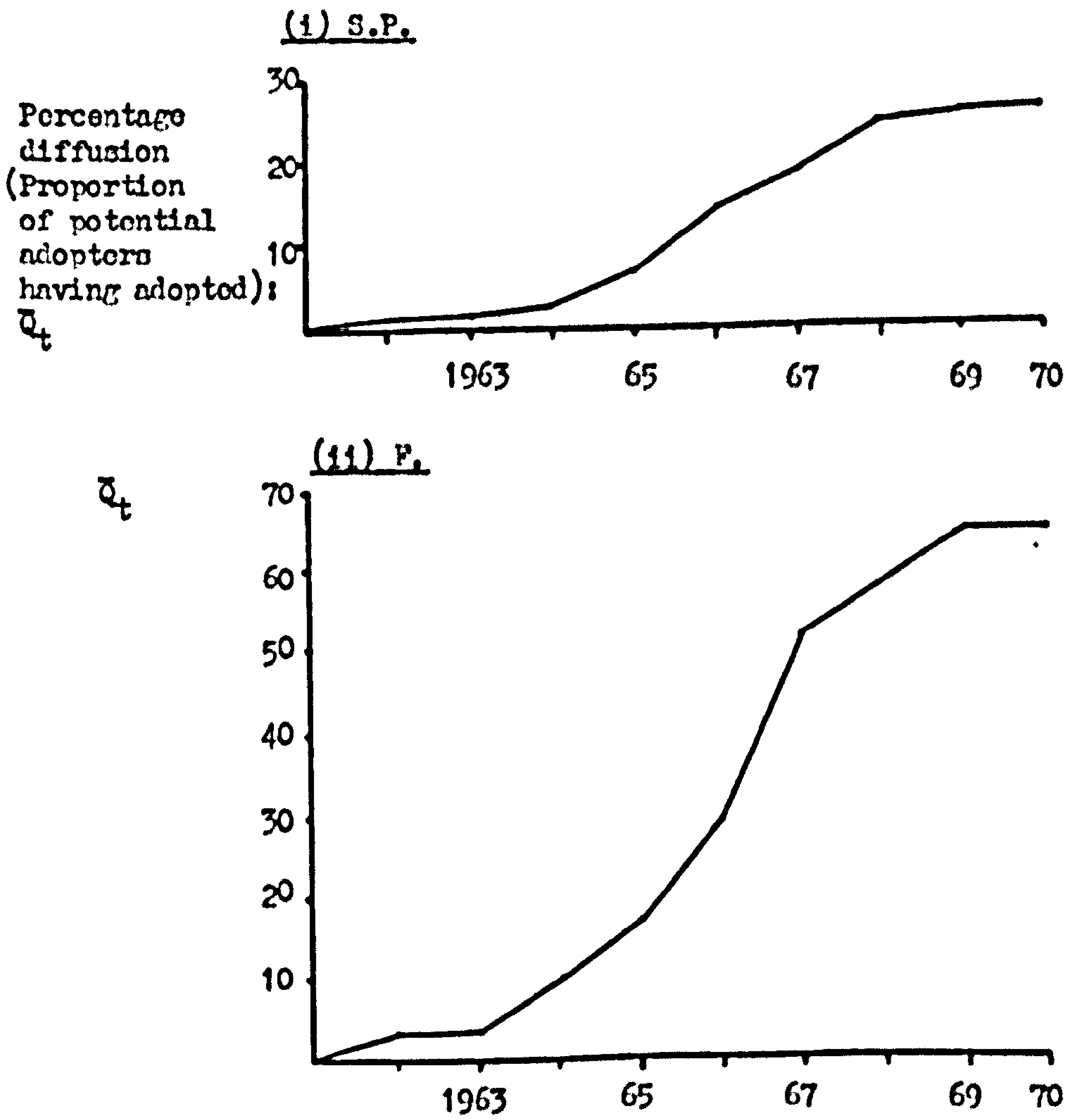
Seven adopters in the sample therefore implies 13.5 adopters in the industry.

In general conclusion then, one must say that for one or two cases, the sample may not be very representative, with the implication for time series regressions that, although year-to-year changes in diffusion may be quite accurate, the overall level might be less so. On the other hand, none of the samples seem too poor to immediately invalidate their use in this thesis;

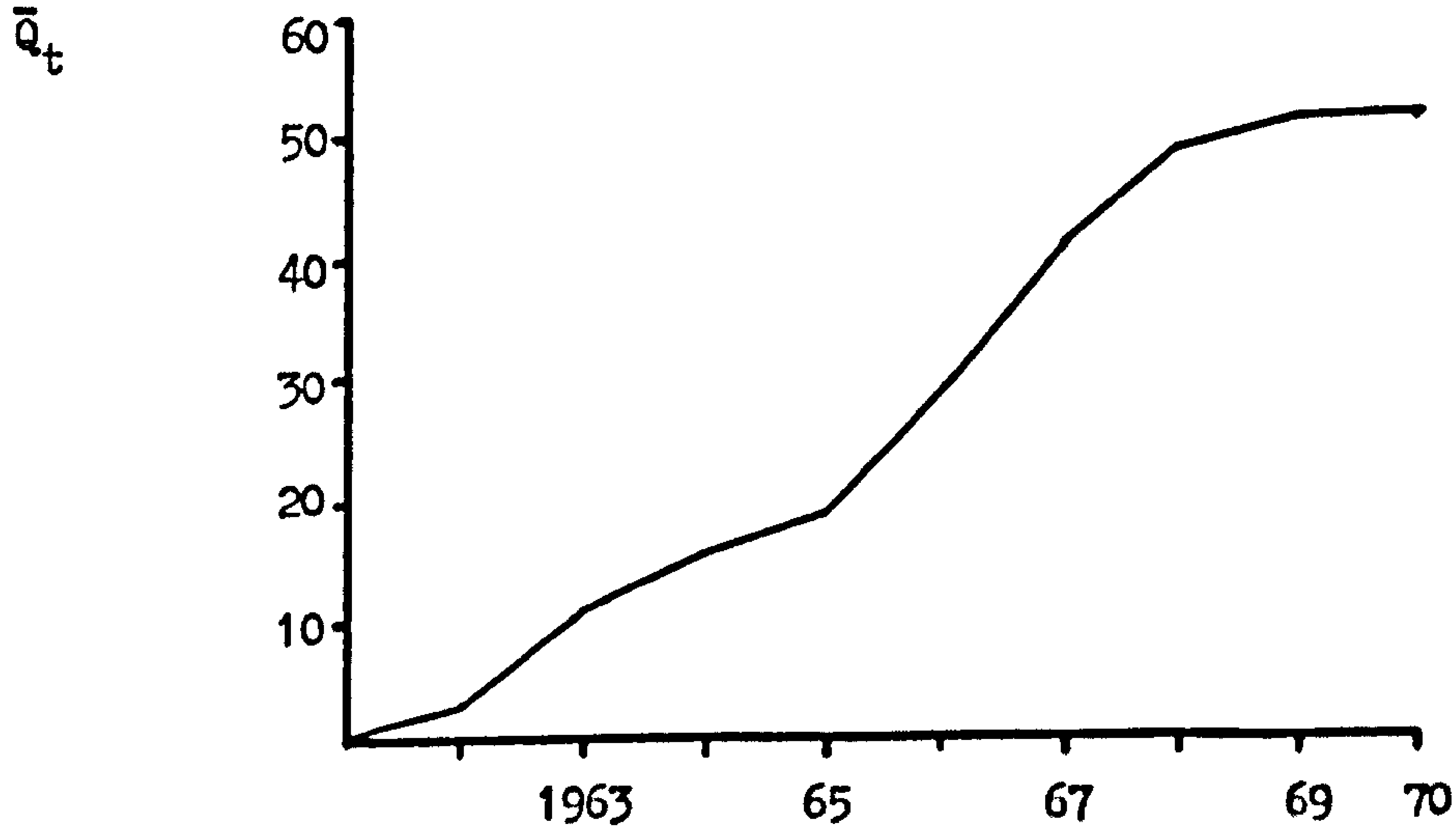
needless to say, random samples are a practical impossibility and it would be difficult to envisage achieving much higher responses than did the N.I.E.S.R.

4. <u>The following figures present the time series data</u> on percentage diffusion, $\bar{Q}_t$, used for chapter 6. ($\bar{Q}_t = (m/n)_{jt}$ where m_j = the number of adopters in the jth industry; n_j = the total feasible set of adopters.) Section 2 of this Appendix has included the definitions of and sources for n, Section 3 has described how m_t has been computed from the N.I.E.S.R. samples and, for those innovations for which data on m_t was obtained from other sources, these sources are given under the relevant figures below.

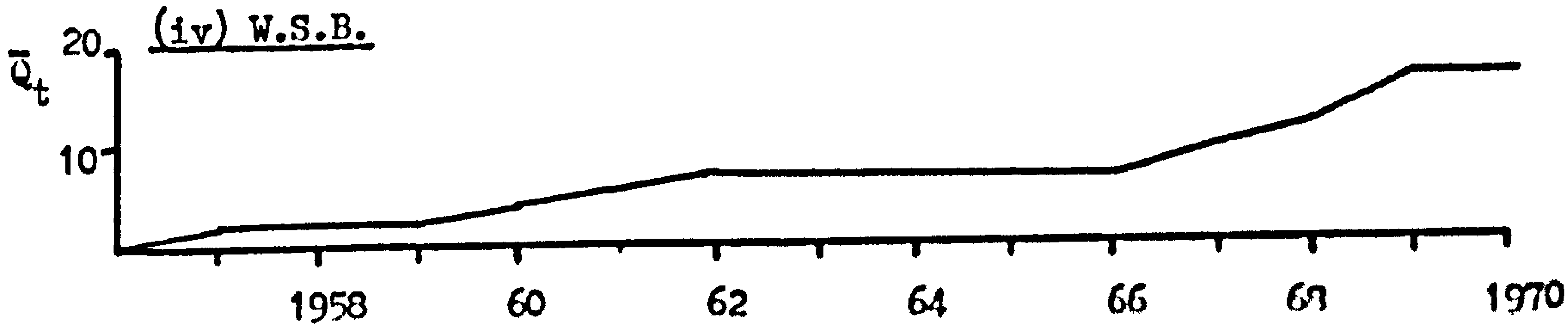
<u>Figure A.2.1: Time series for the percentage of potential adopters having adopted.</u>



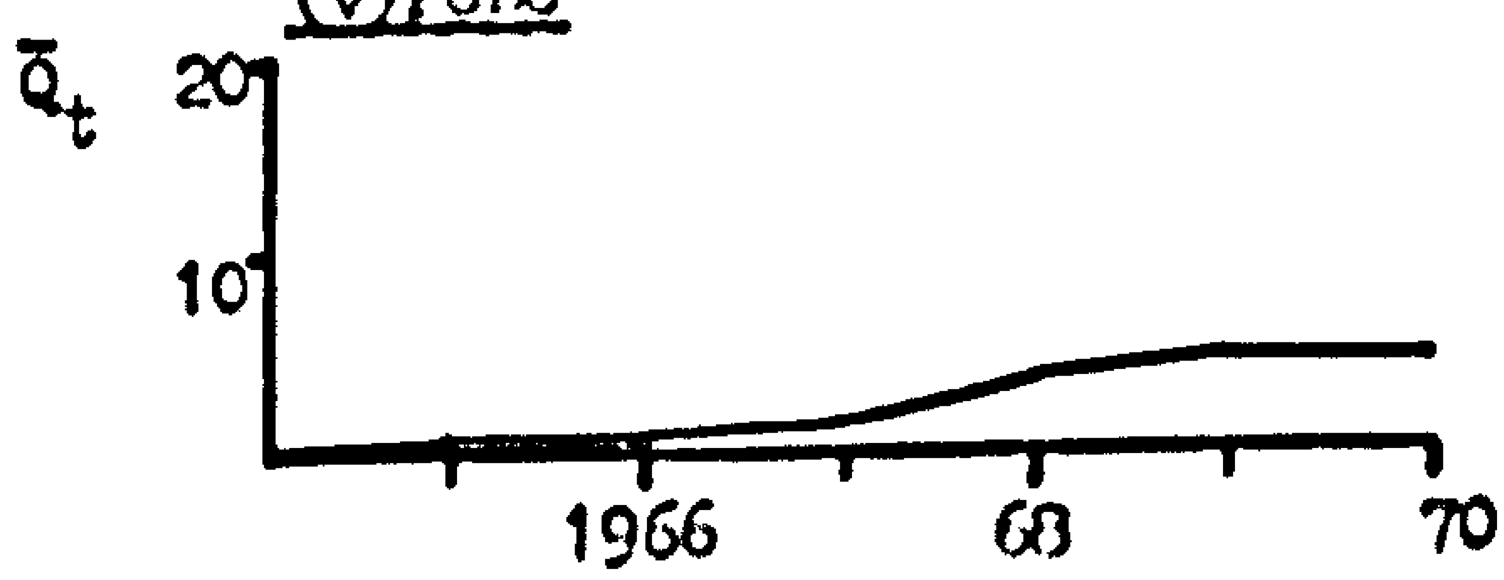
(iii) S.F.



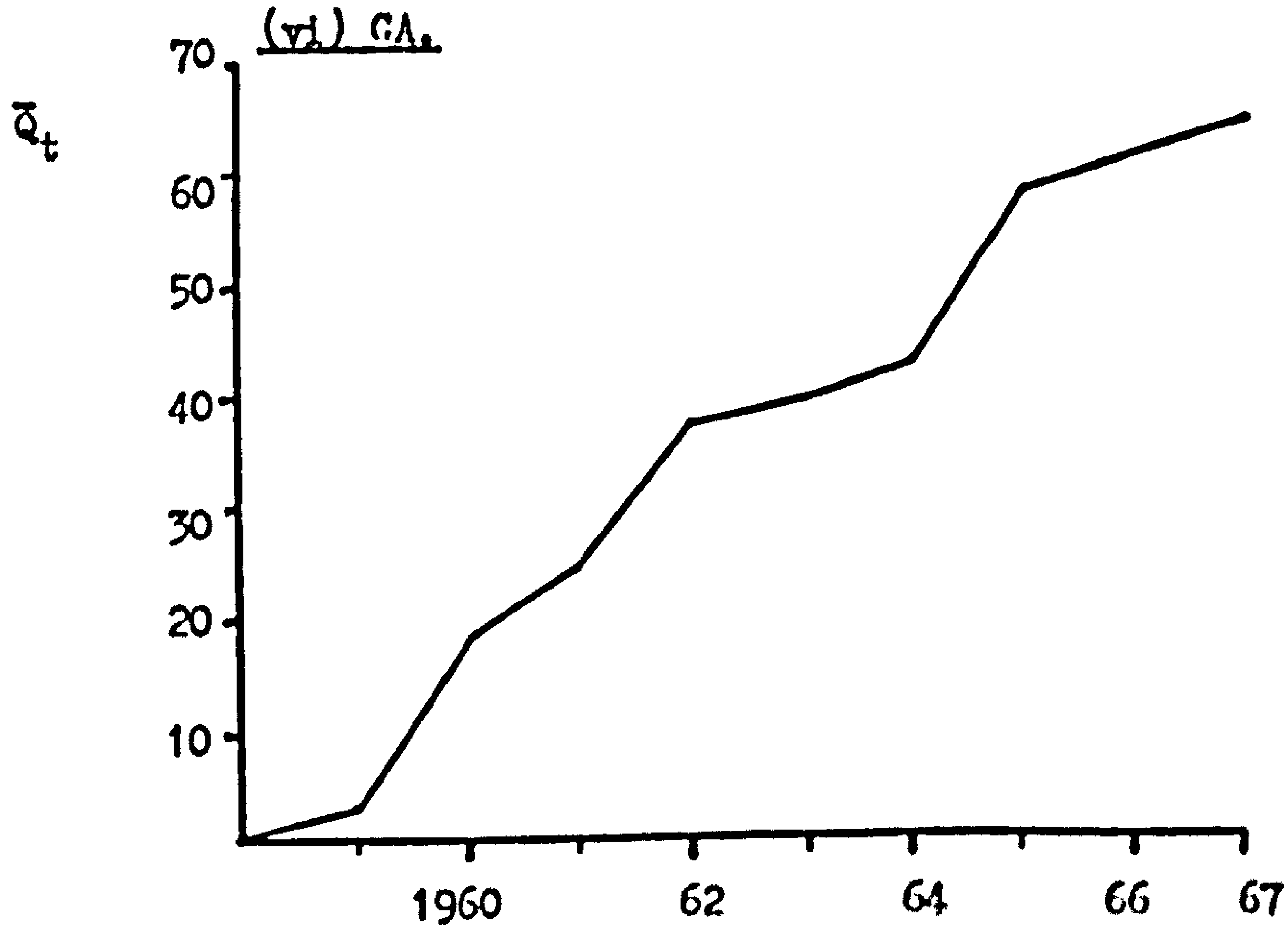
(iv) W.S.B.



(v) PCBC[1]

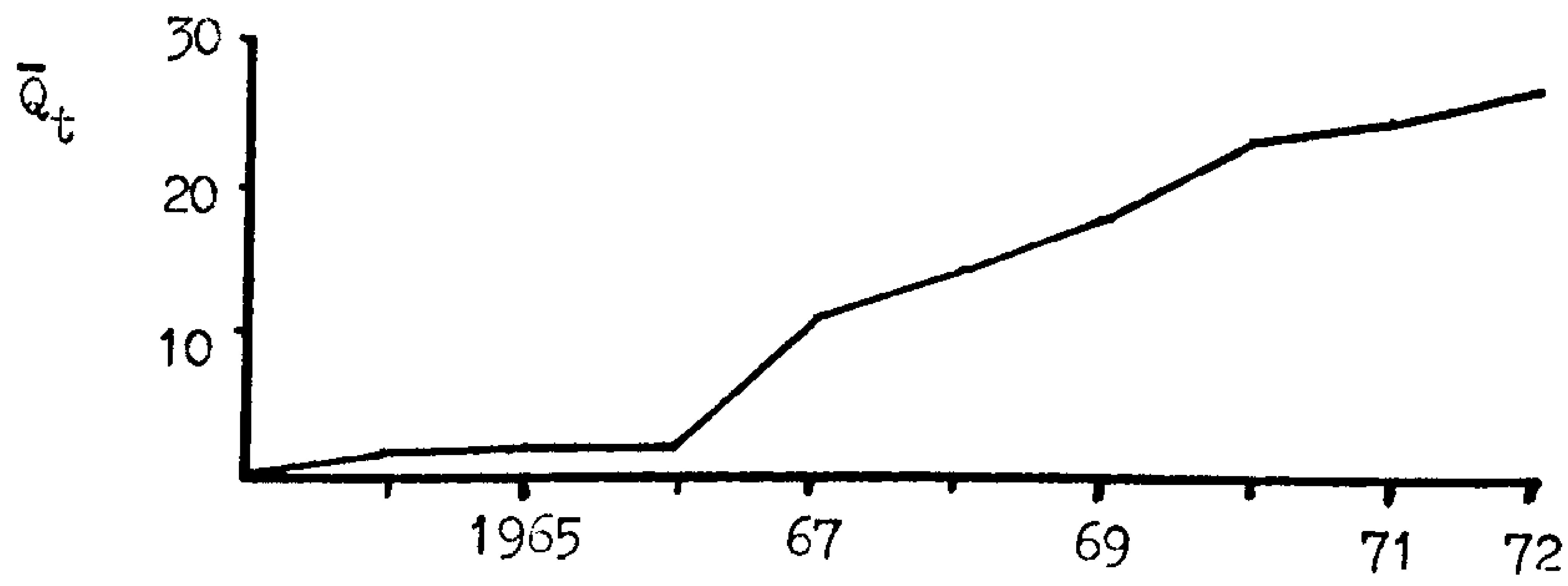


(vi) CA.

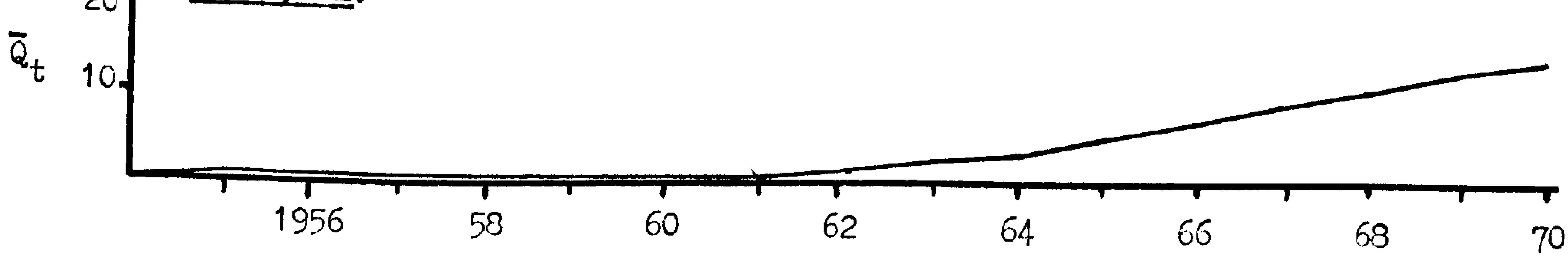


1. Source: various editions of "Computer Survey".

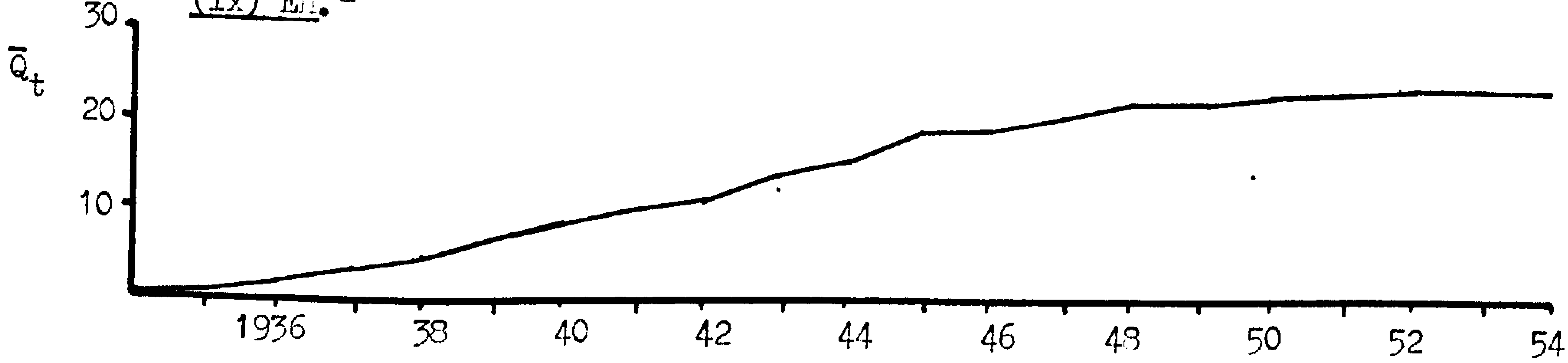
(vii) CT.[1]



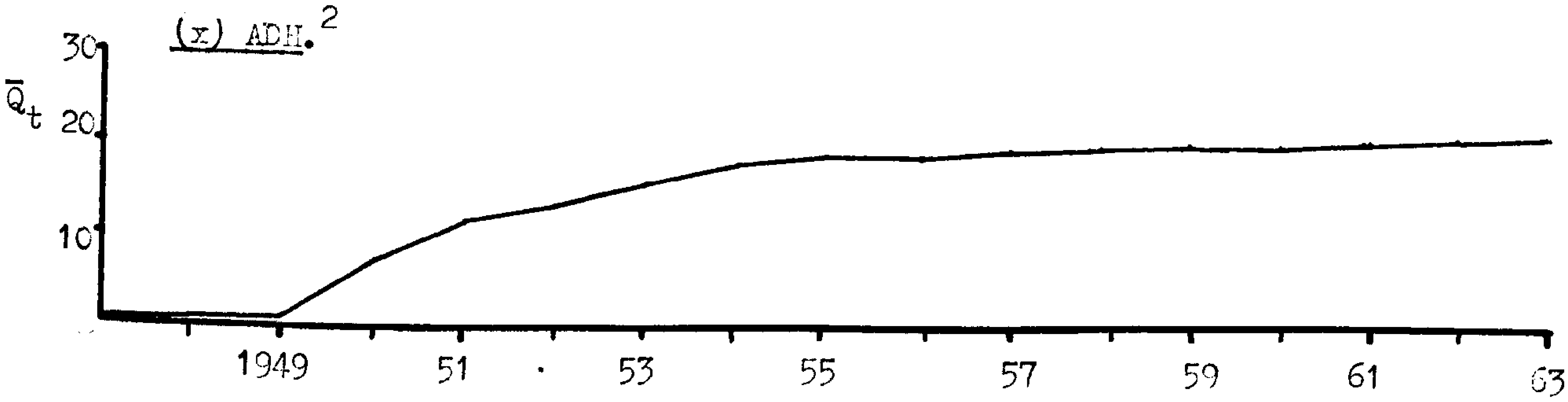
(viii) SL.



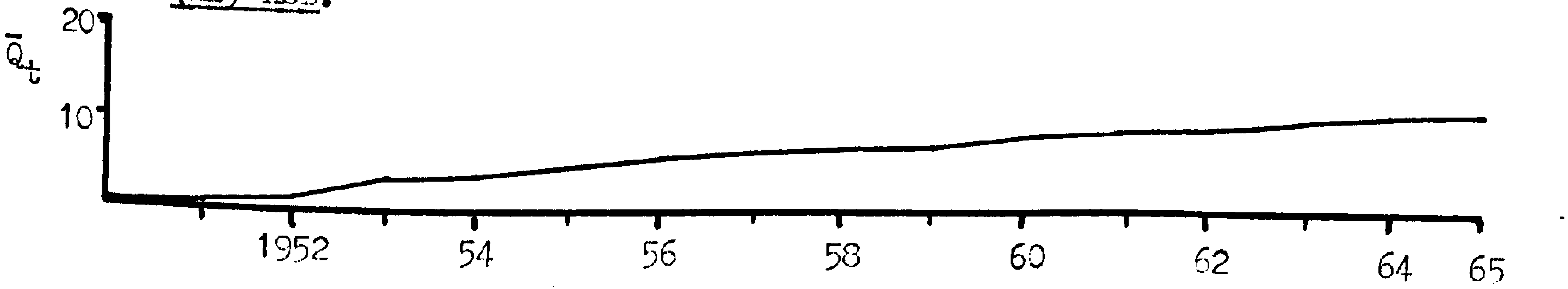
(ix) EH.[2]



(x) ADH.[2]



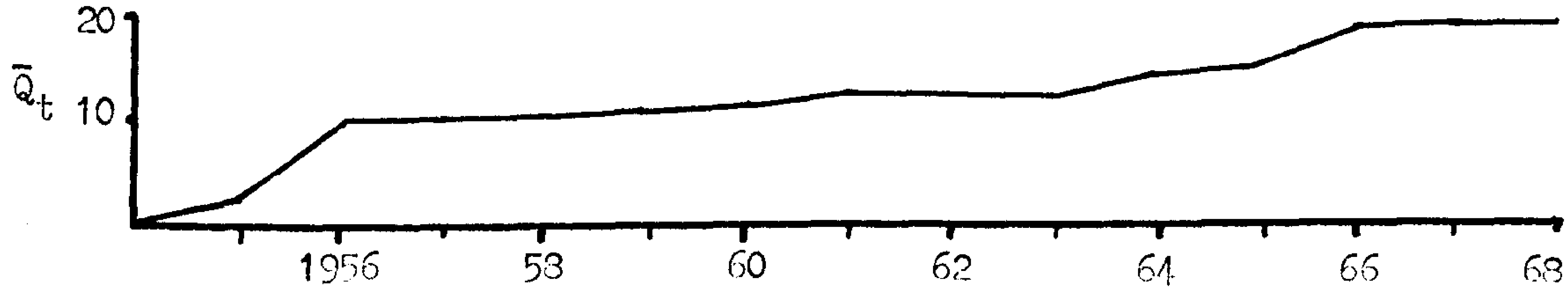
(xi) ASB.[2]



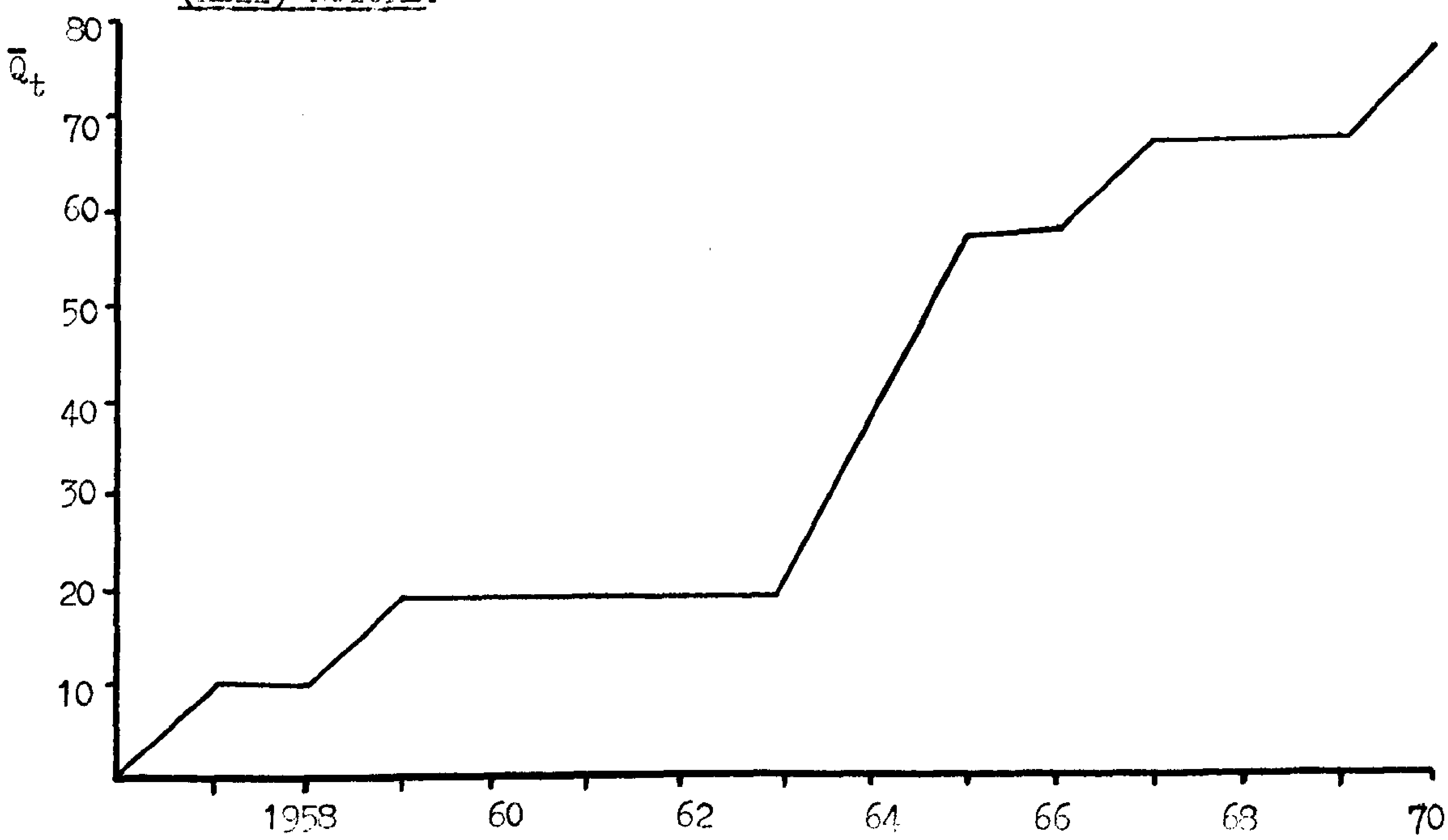
1. Source: various editions of "Computer Survey"

2. Source: Metcalfe, (1970) op.cit.

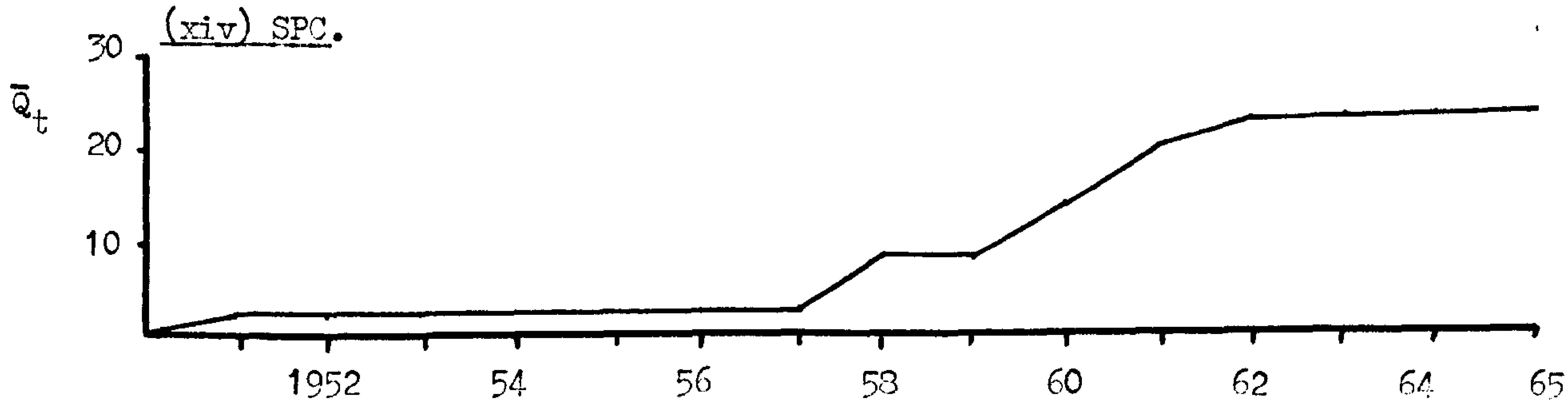
(xii) TC.[1]



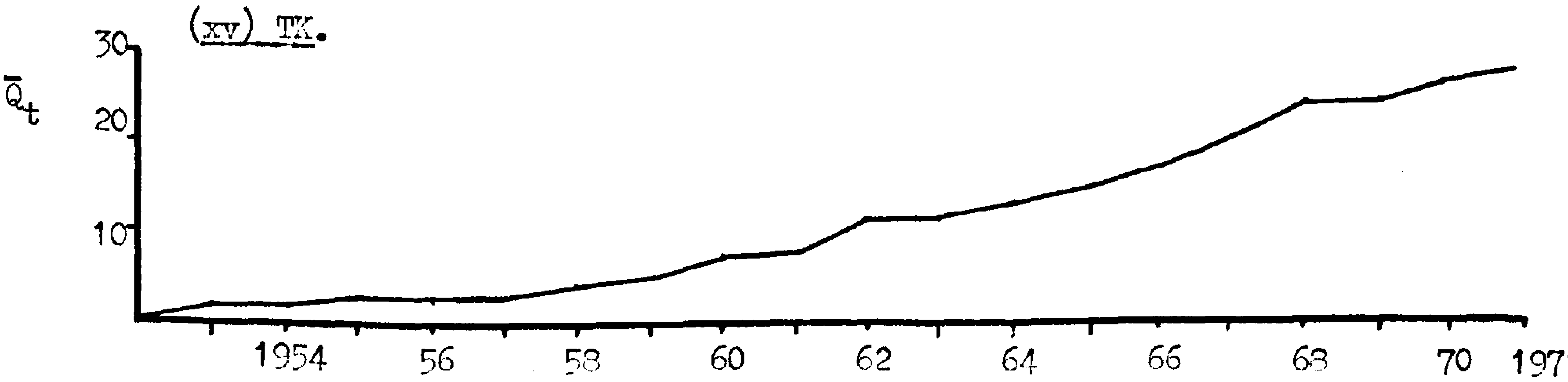
(xiii) NCTURB.



(xiv) SPC.

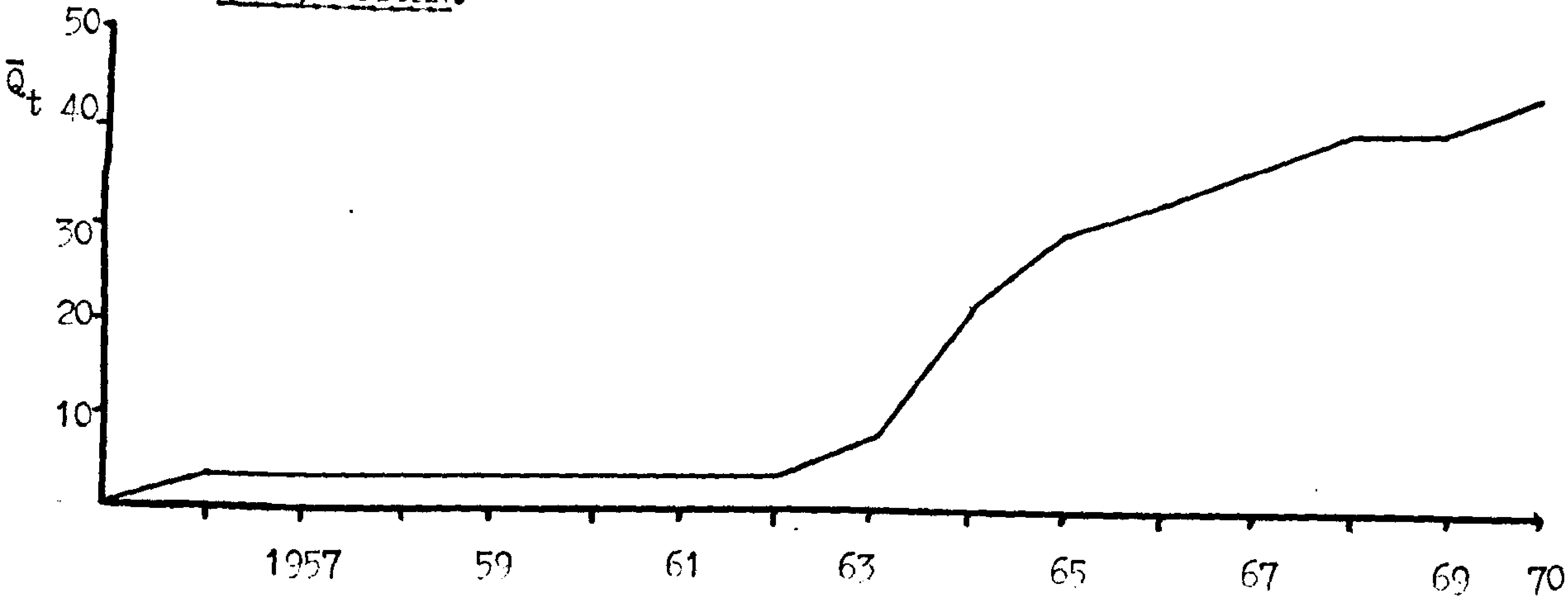


(xv) TK.

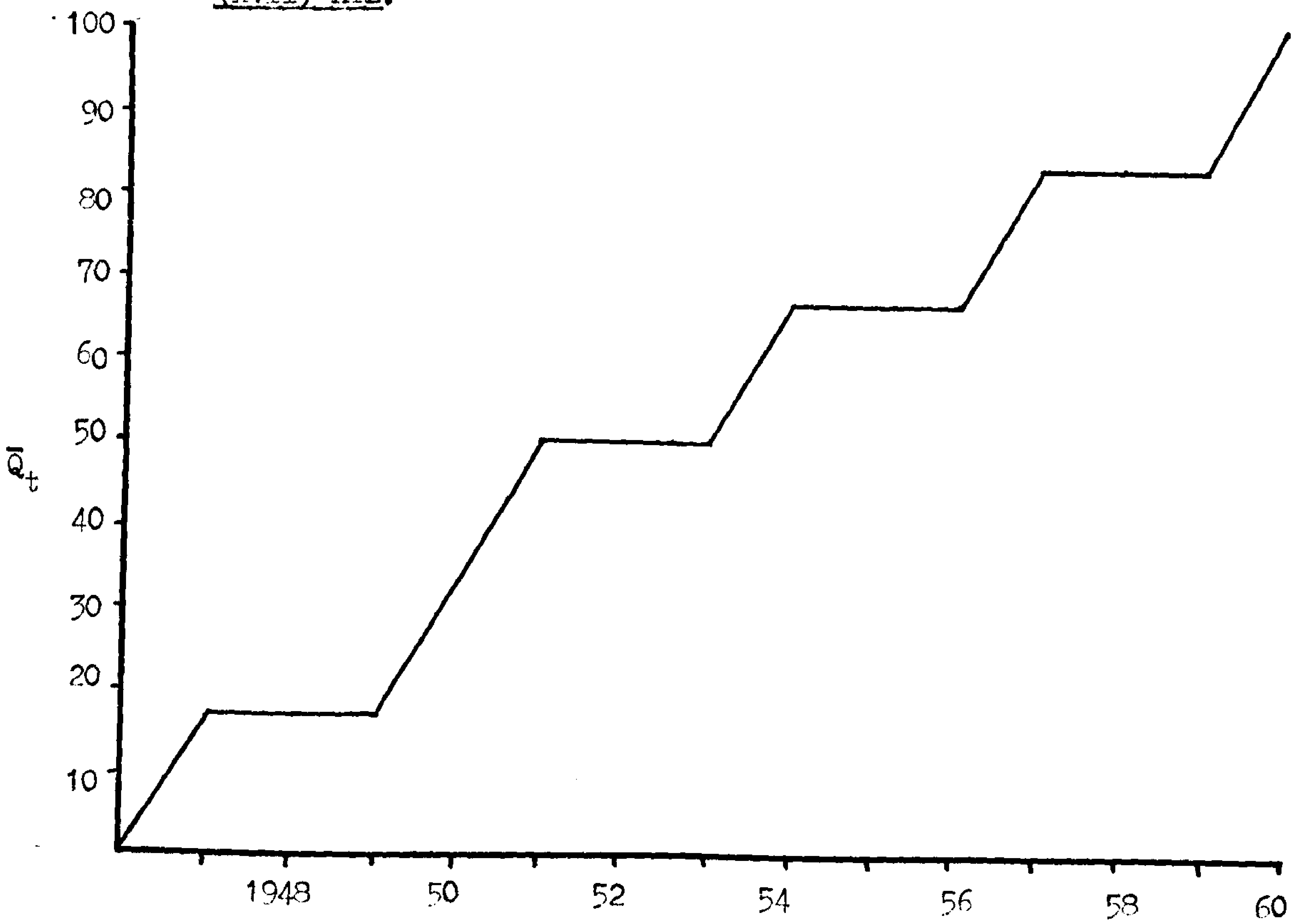


1. Source: T.W.K. Scott (op.cit.)

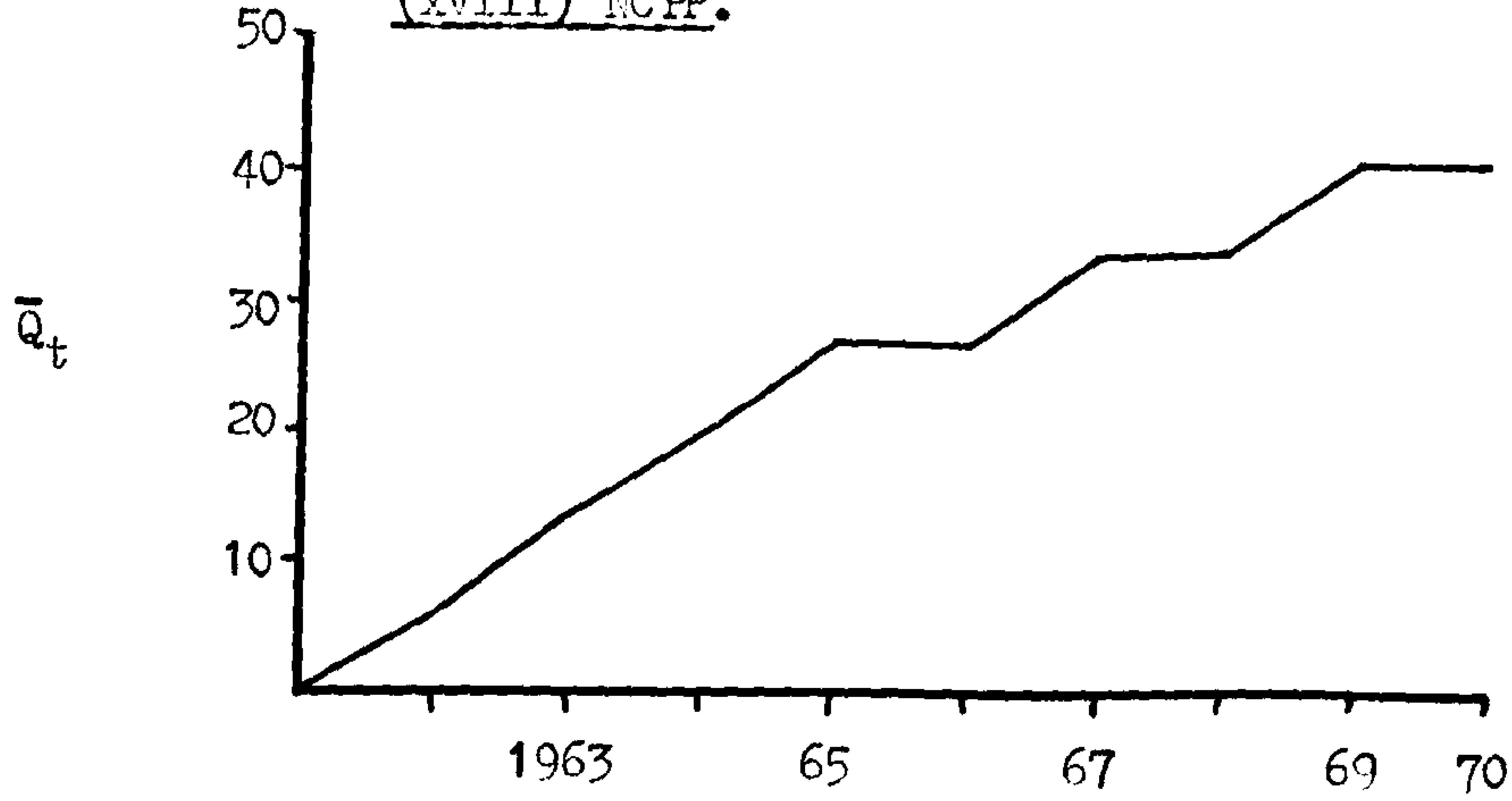
(xvi) NCTURN.

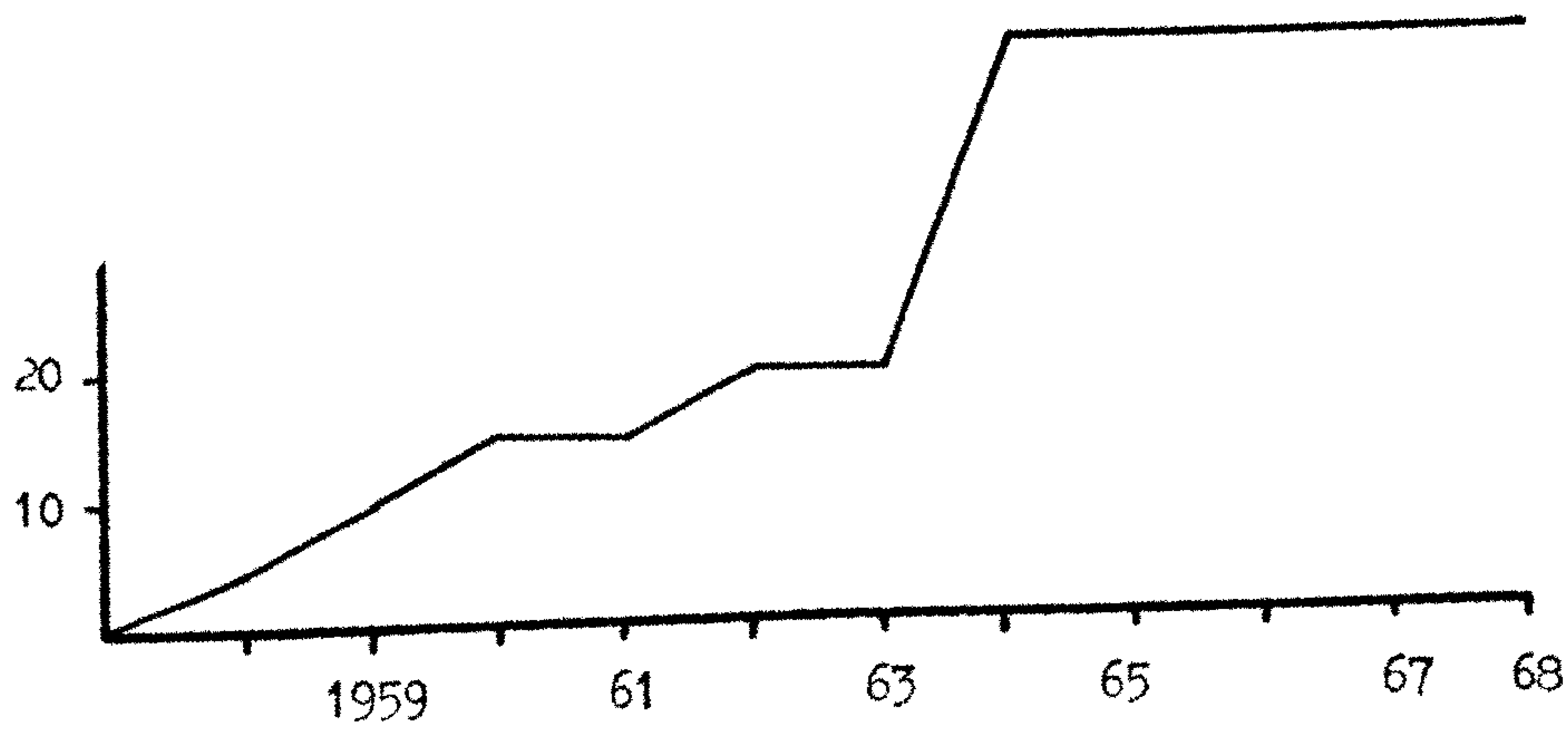


(xvii) ATL.

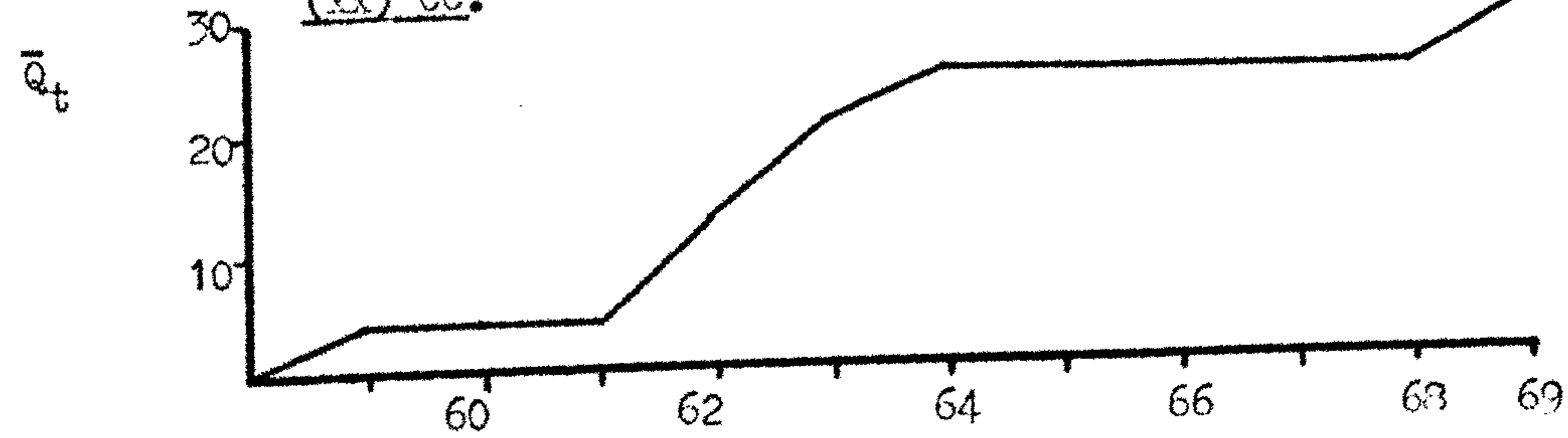


(xviii) NCPP.

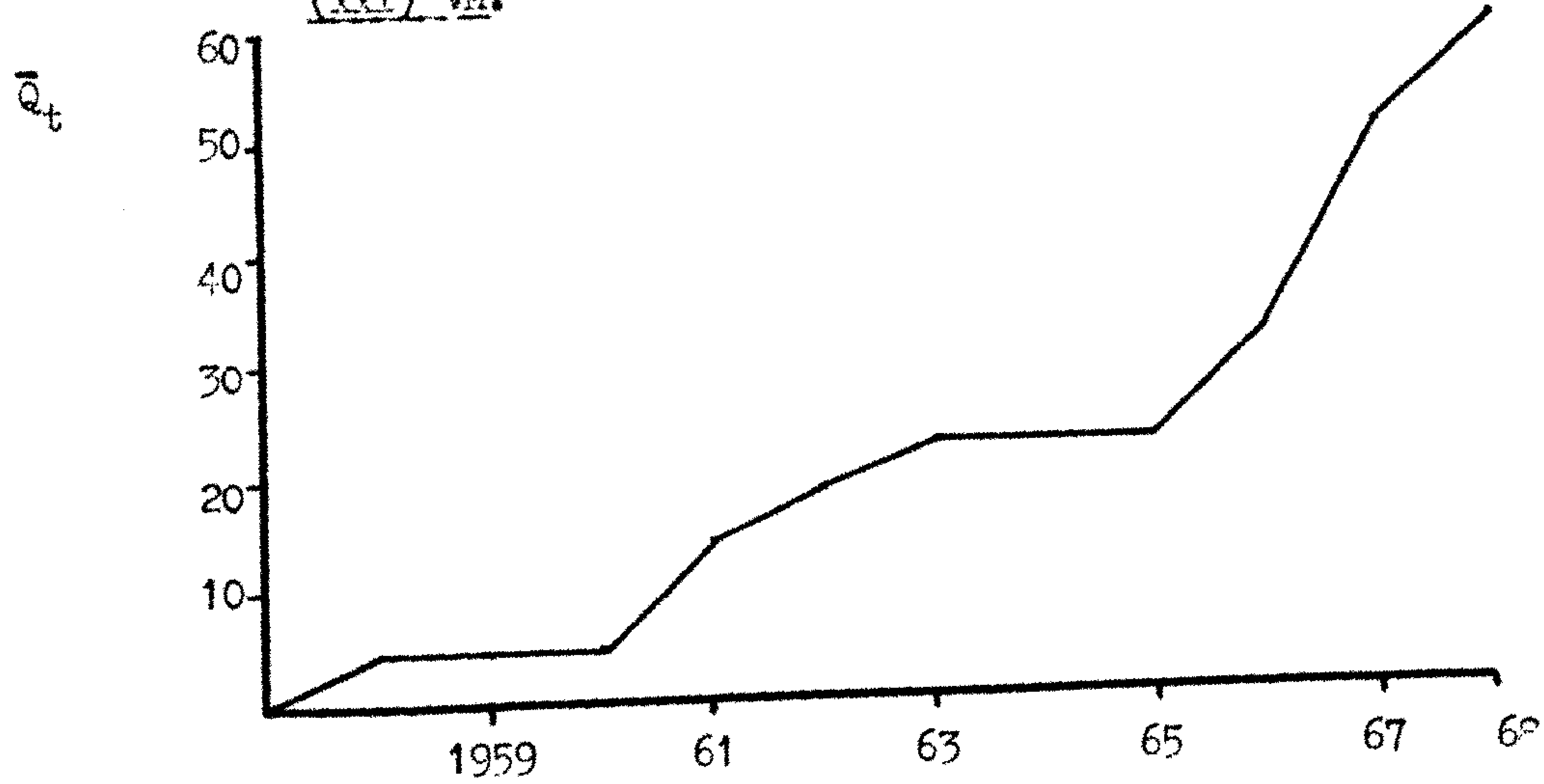




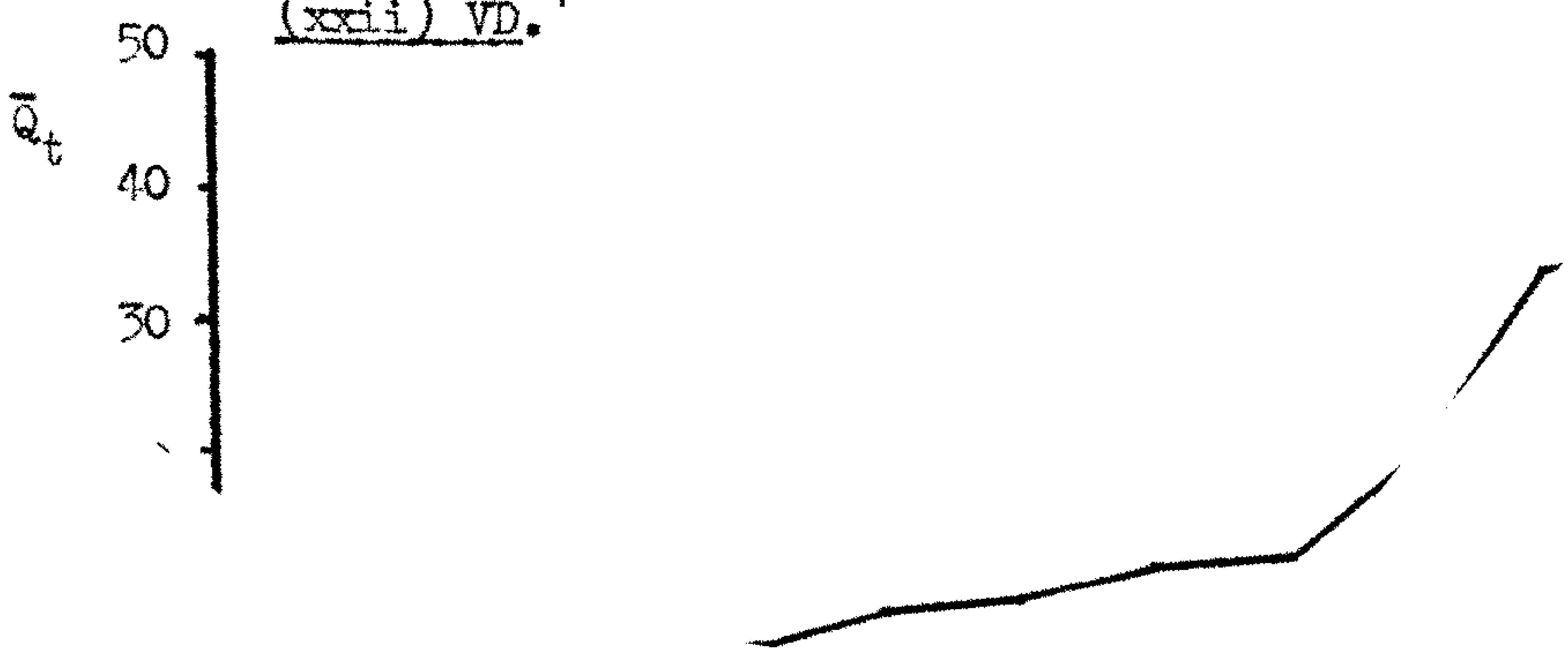
(xx) CC.[2]



(xxi) VM.[3]



(xxii) VD.[4]



Sources for (xix) - (xxii):

1. various editions of the B.I.S.F. Annual Statistics.
2. U.N.E.C.E. report (1968) op.cit.
3. Barraclough, J.I.S.I. (1969), op.cit. and various editions of the B.I.S.F. Annual Statistics.
4. Flux, J.I.S.I. (1965), op.cit. and various editions of the B.I.S.F. Annual statistics.

Appendix 3. The sources and quality of the data used to measure C_{tj} in the time series regressions of Chapter 6.

Although the results reported in Chapter 6 section 6 including a cyclical variations variable refer only to the use of capacity usage as the operative measure, three alternative measures were also used with little success. In this appendix the exact specificationsof these measures are discussed and problems of measurement outlined.

(i) D_{tj} was used to represent the level of industry j's demand in year t; to standardise for differences in industry demand, it was expressed as an index with a base of demand in the year of introduction of the innovation concerned.

Two major problems of measurement presented themselves. For some industries, data was just not available at the level of disaggregation required and so more broad industry definitions had to be used e.g. for the printing press industry. Furthermore, sometimes data was not available for demand at all; fortunately output data was available in these cases and was used as an (admittedly inferior) substitute.

D_{tj} is defined as a real variable, therefore constant price or volume-based series were always used.

The first column of table A.3.1. lists the operational measures used and the notes to that table present the sources for these series.

(ii) INV_{jt} was used to represent the level of gross investment in all plant and equipment in industry j in year t. It was standardised across industries by measuring it as an index based on investment in the start year. Again there are two empirical problems, both stemming from the scarcity of time series disaggregated data on investment. There have been small annual Censuses of Production for most of the post-war years and these provided series for gross investment at current prices at various and variable industry aggregations. [1]

1. Most of which are summarised in the Board of Trade Journals of 27.12.1963; 20.12.1968 and 30.12.1970, although data for some years in the 50's is only available in the original reports.

Unfortunately, more often than not these aggregations are higher than those required for my purposes; moreover satisfactory price deflators for these series are generally unavailable. The second column of table A.1.3. lists the series which were, in the event, used; and, as can be seen, in the worst cases, the required and actual level of aggregations for investment data (such as 'All drink industries' for 'malting') diverge by so much as to render use of the data rather questionable. Price deflators for investment were calculated from data published in the 'Monthly digest of Statistics' on investment in current and constant prices for the eight broad industry classifications which encompass all of the industries within my sample. To indicate the level of imprecision that their use leads to, a typical example might be given: the series for investment by the crude steel sector of the Iron and Steel Industry is calculated from data on investment in current prices by all of Iron and Steel, which is then deflated by a derived price index of investment goods used by the broad 'Metal Manufacture' group of industries. Such approximations obviously pose serious question marks against the validity or usefulness of INV_{tj}, which are confirmed by its poor performance in the empirical work.

(iii) U_{tj} was used to represent the level of unemployment in industry j in year t. More specifically, it was measured as the ratio of wholly unemployed male and female workers to the number of employees in employment in the industry [1]. The only problem encountered (apart from the well-known unreliability of unemployment figures at the industry-level) was the usual one of not having access to sufficiently disaggregated data; however this was less frequent in this case as the industrial breakdown of unemployment figures is quite fine. Column 3 of table A.3.1. lists the series actually used.

1. Data for both of these series is presented in the Department of Employment Gazette relating to the second week in November of each year.

(iv) C_{tj} was used to represent the level of capacity usage in industry j in year t. This is the one measure for which results are spelled out in detail in chapter 6. As already explained, it was computed as the ratio of actual output to capacity output in each year. Series for the latter were calculated using the most simple Wharton school technique of interpolating between peaks of output. Unfortunately it has proved impossible to obtain data on output for many industries and in those cases, demand has been used as a surrogate. In fact, for only four industries was data available for both demand and production, therefore, with the following exceptions the same data series have been used to compute both D_{tj} and C_{tj}:

SP, SF, F, WSB, PCBC	(production of all paper and board.)
BOP, CC	(crude steel output.)
VD	(total steel output.)
VM	(special steels output.)

For all other industries, the series used are listed in the first column of table A3.1.

As for D_{tj}, INV_{tj} and U_{tj} there is the problem in some cases of the data being at a more aggregated level than is desirable, given the fine industry classifications required in this study.

Table A.3.1. Data series used to represent economic activity variables.

Innovation	D_{tj}, C_{tj} Measure of demand/ production	INV_{tj} Investment series[10]	U_{tj} Unemployment series[11]
SP; SF; F; WSB; PCBC	apparent consumption of all paper and board[1]	Paper and Board (MLH 481)	Paper and Board
BOP; CC	crude steel deliveries[2]	Iron and Steel (MLH 311 - 13)	Iron and Steel (general)
VD	total steel deliveries[2]	Iron and Steel (MLH 311 - 13)	" "
VM	special steel deliveries[2]	" "	" "
EH; ADH; ASB	total yarn sized [3]	Cotton, man-made fibres etc., textiles (MLH 412, 413 & 423)	cotton weaving
NCTURN	deliveries of machine tools[4]	Mechanical engineering (MLH 331-9, 341, 342 and 349)	Machine tools
NCTURB	deliveries of turbines[5]	" " "	Industrial engines
NCPP	index of production for mechanical engineering[6]	" " "	Other non-electrical engineering
GA	beer production [7]	Drink industries	Brewing and malting
SL	woven cloth production[8]	as for EH etc.	as for EH etc.
TK	deliveries of clay bricks[9]	Bricks and cement (MLH 461-4 and 469)	Bricks and fireclay goods
SPC	index of production for shipbuilding[6]	Shipbuilding and marine engineering (MLH 370)	Shipbuilding
CT	consumers' expenditure on newspapers(constant prices) [7]	Paper products, printing and publishing	Printing and publishing of newspapers and periodicals
TC	deliveries of all carpets [8]	Other textile industries	Carpets
ATL	production of cars [9]	Motor Vehicles (MLH 381)	Motor Vehicles

Sources

1. B.P.B.M.A. Reference tables 1970, 71, 72.
2. Annual Statistics of the Iron and Steel Board and B.I.S.F.
3. Metcalfe's M.Sc. thesis.
4. Business Monitor on Metalworking machine tools.
5. " " Industrial engines.
6. Economic trends, November 1972.
7. " " various issues.
8. Textile Council's Quarterly Review.
9. Housing Statistics, various issues.
10. 'Trade and Industry' / 'Board of Trade Journal'; 27.12.63; 20.12.68; 30.12.70. and various Census of Production reports and Monthly Digest of Statistics (various).
11. Department of Employment / Ministry of Labour Gazette various issues.

Appendix 4 : The data used to estimate the Quasi-Engel curves.

This appendix reproduces the data used to estimate the Quasi-Engel curves reported in Chapter 7.

For each innovation and industry [1], firms are grouped into a number of size ranges. For each range the tables show the number of firms in that range, the penetration of the innovation (i.e. the proportion of firms having adopted) and the average size of firm (measured in numbers employed.)

The date for which these measurements apply is shown in brackets after the name of the innovation. As reported in chapter 7, there was very little discretion in the choice of year for which the curves were calculated. For some innovations the only data available was for one year and for given size ranges. Whilst for others penetration data was available for a number of alternative years, information on firm size was not. Moreover these curves can only be estimated so long as penetration does not reach 100% in any range - the normit of 100% is, of course, infinity - therefore in some cases, the choice of estimating year was further restricted.

For most of the innovations, the sources for these data are given in Appendix 2; for one or two, however, additional sources are indicated in footnotes to the tables.

At the end of this appendix the method for calculating $(\widehat{\sigma/\beta})$ for ATL is described. Estimation of an Engel curve for the car industry was impossible as, for these purposes, it is defined as including only 6 firms: obviously no meaningful groupings are possible for such a small number.

1. For a number of industries, data is only available for a sample of firms - the samples are usually identical to those outlined in Appendix 2.

Table A4.1. The data used in tabular form.

PAPER AND BOARD

(i) Special presses (1970)

Size range		Penetration	Average employment
> 50,000 tons*	(14)[1]	.786	4195
30 – 50,000 "	(10)	.6	616
20 – 30,000 "	(8)	.5	424
10 – 20,000 "	(9)	.333	246
0 – 10,000 "	(4)	.0	64

(ii) Foils (1966)[2]

> 50,000 tons	(7)	.857	4195
20 – 50,000 "	(10)	.500	531
5 – 20,000 "	(7)	.2857	183
0 – 5,000 "	(0)	?	38

(iii) Synthetic Fabrics (1966)[2]

> 50,000 tons	(7)	.857	4195
20 – 50,000 "	(10)	.400	531
5 – 20,000 "	(7)	.143	183
0 – 5,000 "	(0)	?	33

(iv) Wet Suction boxes (1970)

> 50,000 tons	(7)	.429	4195
20 – 50,000 "	(10)	.300	531
5 – 20,000 "	(7)	.143	183
0 – 5,000 "	(0)	?	33

(v) Process control by computor (1970)

> 50,000 tons	(14)	.2143	4195
20 – 50,000 "	(18)	.0556	531
0 – 20,000 "	(66)	.0152	106

* (tons of paper produced)

IRON AND STEEL

(i) Basic Oxygen Steelmaking (1968)

Range of employees		Penetration	Average employment
8500 +	(6)	.833	18300
4000 – 8499	(6)	.5	7247
1 – 3999	(8)	.125	2298

(ii) Continuous Casting (1969)

8100 +	(8)	.375	15800
3000 – 8099	(8)	.375	5670
1 – 2999	(8)	.125	1375

1. brackets after size range denotes number of firms in sample within that range.

2. After 1966, penetration in the >50,000 tons group reached 100%. As the normit for 100% is infinity,it would have been technically impossible to fit the line

(iii) Vacuum degassing (1968)

Employment Range		Penetration	Average Employment
8400 +	(8)	.625	15780
6000 - 8399	(7)	.571	7007
2000 - 5999	(8)	.500	3272
1250 - 1999	(7)	.429	1635
1 - 1249	(8)	.125	529

(iv) Vacuum Melting (1968)

1000 +	(8)	.875	3062
350 - 999	(7)	.571	496
1 - 349	(7)	.285	207

LANCASHIRE WEAVING

(i) Electric Hygrometers (1956)

Size Range *	Penetration	Average Employment
> 800 looms	.82	785.5
401 - 800 "	.29	271.8
201 - 400 "	.09	135.9
1 - 200 "	.007	45.3

Source (for all three innovations): Metcalfe p. 152 and footnote p. 153, (1970- op.cit.) Number of firms in each range, not given by Metcalfe.

(ii) Accelerated Drying Hoods (1956)

> 800 looms	.593	787.5
401 - 800 "	.25	271.8
201 - 400 "	.11	135.9
1 - 200 "	.0035	45.3

(iii) Automatic Size Boxes (1956)

> 2000 looms	.19	1332.7
801 - 2000 "	.11	634.2
401 - 800 "	.06	271.8
1 - 400 "	.00	71.0

* Number of looms employed

NUMERICAL CONTROL

(i) Printing press industry (1970)

Employment Range		Penetration	Average Employment
500 +	(9)	.667	1338
76 - 499	(7)	.429	199
1 - 75	(8)	.125	39

(ii) Turning machine tools industry (1970)[1]

500 +	(12)	.917	1930
200 - 499	(13)	.461	293
1 - 199	(11)	.0909	71

1. The classes had to be larger than average due to the need to avoid 100% penetration in the top class.

(iii) Turbine industry (1970)

Employment Range		Penetration	Average Employment
1000 +	(7)	.857	3786
1 - 999	(2)	.50	275

MALTING : Gibberellic Acid (1967)

Employment Range		Penetration	Average Employment
70 +	(8)	.875	168.5
25 - 69	(10)	.700	68.9
13.- 24	(6)	.500	15.9
1 - 12	(7)	.571	5.9

WEAVING: Shuttleless Looms (1970)

Employment Range		Penetration	Average Employment
2000 +	(7)	.857	2951
500 - 1999	(8)	.75	1199
175 - 499	(8)	.50	274
125 - 175	(8)	.25	136
1 - 124	(9)	.11	61

BRICKS : Tunnel Kilns (1971)

Employment Range		Penetration	Average Employment
200 +	(13)	.923	559
50 - 199	(14)	.286	119
30 - 49	(14)	.286	39
1 - 29	(15)	.067	21

SHIPBUILDING: New methods of Steel plate cutting (1961)

Employment Range		Penetration *	Average Employment
3000 +	(5)	.8	4850
500 - 2999	(6)	.167	1526
1 - 499	(3)	.0	265

* given problems of non-response in the original sample, these proportions are tentative.

PROVINCIAL NEWSPAPERS: Computer Typesetting (1972)

Employment Range		Penetration	Average Employment
700 +	(12)	.417	1008
400 - 699	(10)	.400	493
300 - 399	(11)	.364	323
240 - 299	(12)	.085	266
1 - 239	(12)	.167	214

CARPET INDUSTRY : Tufted Carpet machines (1966) *

Employment Range	Penetration	Average Employment
500 +	.80	1393
200 - 499	.71	326
25 - 199	.40	95
1 - 24	.0	16

* This table is based on a sample of 1/3rd of all firms in the industry by T.W.K. Scott (op.cit.) New entrants are excluded.

CAR INDUSTRY : Automatic Track Lines.

Whilst it is impossible to compute a Quasi-Engel curve of any meaning for this industry of only 6 firms, by happy coincidence an estimate of (σ/β) is still possible. ATL is the one innovation in the sample which had diffused 100% at the time of study and, consequently, an alternative approach to (σ/β) is available.

In this unique case, observations are available for all firms on the date of adoption (d_i). Therefore it is possible to use the relationship indicated in Appendix 3 to Chapter 5. As ATL is a group B innovation, the following implication can be derived from the model (equation 5. A3.6):

$$d_i = -\frac{1}{\Psi}\log\alpha - \frac{\beta}{\Psi}\log S_i - \frac{1}{\Psi}\log\epsilon_i$$

(table 6.6.1. indicates that $\Omega = 0$ for ATL)

On regressing d_i against $\log S_i$ for the 6 firms in this industry, $\widehat{(\beta/\Psi)} = .395$.

Furthermore, from table 6.4.1., the estimate of b for ATL, using the cumulative normal, is .882

As $b = \Psi\left(\beta^2\sigma_s^2 + \sigma^2\right)^{-1/2} = (\Psi/\beta)\left(\sigma_s^2 + \sigma^2/\beta^2\right)^{-1/2} = .882$

Therefore, $\sigma_s^2 + (\sigma/\beta)^2 = (.882 \times .395)^{-2}$

and as σ_s^2 is given in Appendix 5 as .3645, $\sigma^2/\beta^2 = 8.26 - .3645 = 7.8955$

In passing, it is interesting to note the direct link between the speed of diffusion, the slope of the Engel curve and the size elasticity in the 'time-lag before adoption' equation used so much in previous research (see section 3 of chapter 2.)

Appendix 5. The Sample industry firm size distributions.

In section 4 of chapter 5, firm size is assumed to be lognormally distributed within each industry. This appendix provides a test of this assumption, using the size distributions actually observed for the sample industries and in doing so generates estimates of the parameters of these empirical distributions.

As a preliminary, a brief survey is presented of the theoretical assumptions needed for lognormality, the empirical validity of such assumptions and of the techniquesused in the past to test empirical size distributions for lognormality. Then, in a second section, the lognormal assumption is tested for the sample industries and estimates of parameters are computed in each case.

1. A survey of past work.

The lognormal steady state distribution of firm size is generated by a growth process embodied in Gibrat's well-known 'law of proportionate effect'. This states[1] that the growth in firm i's size at time t is given by a random disturbance term ϵ_{it} :

$$\frac{S_{it} - S_{it-1}}{S_{it-1}} = \epsilon_{it} \qquad (A5.1.1)$$

or, alternatively, the size of firm i tends to change over time by a randomly distributed proportion of its size at the start of the time period:

$$S_{it} - S_{it-1} = S_{it-1}\,\epsilon_{it} \qquad (A5.1.2)$$

There are a number of implications of this assumption[2]: all firms in an industry have the same expected growth rate; the dispersion of growth rates is the same for all sizes of firm; the distribution of growth rates is lognormal and the variance of firm size tends to increase over time.

1. See, for instance, Aitchison and Brown (op.cit.) Chapter 3 or R. Gibrat, "Les inégalités économiques," Paris (1931).

2. As noted by P. Hart, 'The size and growth of firms,' Economica, 1962 vol. 24.

Hart[1] tests these assumptions against four sets of data in various U.K. industries, and concludes that they may be only broadly consistent with the facts. His data on firm growth rates is certainly consistent with the first of these implications, a definite conclusion on the second implication is not possible: two out of the four groups considered show no significant differences between large and small firms in the dispersion of growth rates, but one group shows significantly larger dispersion for large firms and another group shows significantly smaller dispersion for large firms. The third and fourth implications are tested against a different set of data and are accepted with some qualifications. Mansfield[2] reports similar findings from data on three major American industries.

In recent years a number of authors have suggested a number of modifications to the model which increase realism at the cost of producing a more complex size distribution. Simon and Bonini,[3] for instance, allow for a constant birthrate (the simple model assumes no births or deaths), this has the effect of producing the more flexible Yule distribution which, interestingly, approximates to the Pareto curve in its upper tail. Saving[4] has introduced the possibility that constant returns to scale (implicit in the basic form of Gibrat's law) only apply over a certain size range - beyond some upper limit there are decreasing returns and below a lower limit there are increasing returns. This has the effect of producing the more complicated four parameter lognormal distribution of firm size.

In order to maintain the simplicity of the model presented in chapter 5,

1. ibid.
2. E. Mansfield, 'Entry, Gibrat's Law, innovation and the growth of firms,' American Economic Review, December 1962.
3. H. Simon and C. Bonini, 'The size distribution of business firms', American Economic Review, September 1958.
4. T. Saving, 'The four parameter lognormal, diseconomies of scale and the size distribution of manufacturing establishments,' International Economic Review, January, 1965.

however, it would be preferable to retain the simple two parameter lognormal assumption. Yet it is clear from the above that the assumptions upon which it is based are likely to be violated to a lesser or greater extent for most industries, and that other statistical distributions might offer a better explanation of the real world.

The obvious question which must be posed is, does the simple lognormal offer an acceptable explanation of empirical size distributions, even though it is based on a number of only approximately valid assumptions? Fortunately, a survey of past work on fitting the lognormal to actual size distributions suggests that, generally, it does.

Hart and Prais[1] in a study of quoted business units in the U.K. for a number of different years in the first part of this century, use two alternative tests of the lognormality assumption. A non-rigorous test, is to plot the cumulative distribution on logarithmic probability paper[2]: a rough indication of lognormality requires that the transformed distribution should describe a straight line. On this purely visual level, Hart and Prais conclude that the lognormal provides a promising fit. As a second, more rigorous test, the parameters of the observed distribution are estimated on the assumption of lognormality, then the goodness of fit is measured using the Fisher test of normality[3] (for log size, of course). In this case, they conclude that there is a slight skewness to the right in the empirical distributions (that is log size has a slightly skewed distribution rather than a symmetrical normal distribution.) However, as the authors note, this may be due to their sample of firms being under-representative of certain size classes in the total population.

1. P.Hart and S.Prais, 'The analysis of business concentration: a statistical approach, 'Journal of the Royal Statistical Society, 1956.
2. See chapter 7, lognormal probability entails measuring log size on the horizontal axis and the normit of the percentage of the population having size below each size level on the vertical axis.
3. This involves calculating the 3rd and 4th moments of the transformed distributions which enables a check of symmetry and kurtosis (convexity.)

As an alternative to the lognormal, the Pareto curve has also been fitted to actual size distributions[1] with interesting results: as a description of overall size distributions, it is clearly inappropriate, however it does provide a close fit for the extreme upper tail for some industries, especially those with relatively few firms. In probably the most comprehensive test to date of the Pareto curve, Quandt[2] fits a number of variations of it, as well as the lognormal, to 32 4 digit SIC industries. Using his own S statistic (probably more rigorous than the standard χ^2 test), he rejects the lognormal as an acceptable distribution in only 6 of the industries, whilst the rejection rate is far higher for each of the four variants of the Pareto distribution which are fitted.

Silbermann[3] suggests another way of testing for lognormality; having estimated the parameters μ and σ^2, he computes the 4, 8, 20 and 50 firm concentration ratio rates predicted. The hypothesis of lognormality is rejected if <u>any</u> of these predicted ratios diverges significantly from the observed ratios. He provides a vast array of different results, an example of which is 82 industries in the U.S. in 1958. Of these, 73 have a 4 firm concentration ratio not significantly different from that predicted by the estimated lognormal parameters.

However, 31 of the 73 do have at least one of the observed 8, 20 or 50 firm ratios significantly different from the predicted values; thus, using this rather stringent test, only just over half the sample may be assumed to have lognormal distributions. However, by concentrating on the upper tail of the distribution, he is testing the Achilles heel of the lognormal and it is still possible that the lognormal may provide an adequate description of the size distributions in their entireties.

1. Simon and Bonini, op.cit., L. Engwall, 'Models of Industrial Structure,' Lexington Books, New York, 1974.
2. R. Quandt, 'On the size distribution of firms,' American Economic Review, June, 1966.
3. I. Silbermann, 'On lognormality as a summary measure of concentration.' American Economic Review, September 1967.

Engwall's[1] results redress the balance somewhat. He uses a variety of test procedures of the lognormal for a wide spread of empirical distributions and claims (rather imprecisely) 'we did obtain fairly good results.' One interesting conclusion of Engwall's is that the lognormal has been found to be most appropriate for 'industries exhibiting moderate change in the number of firms.'

Overall, two points seem incontrovertible. Virtually all observed distributions are positively skewed and may be adequately described by one of the theoretical distributions generated by different versions of the law of proportionate effect; Quandt[2], for instance, finds that only one of 32 industries tested has a distribution not consistent with one of the 6 alternative theoretical distributions (all based on the law). Second, of all the common skewed distributions, the lognormal offers the best overall fit for most industries. However, it must be conceded that there are cases when it does not offer an acceptable explanation. Clearly, some test of the lognormal hypothesis is desirable for the industries included in the present sample, in the light of the above discussion and because of its central role in the development of the model of chapter 5.

2. Lognormality in the sample industries.

It is possible to test the lognormality assumption for all but three of the sample industries.[3]

Figures A5.2.1 show the cumulative distributions on scales consistent with logarithmic probability paper; μ_{sj} and σ^2_{sj} are read off the charts

1. Engwall, op.cit. See also Simon and Bonini (op.cit.) for support for the lognormal.
2. op.cit.
3. For shipbuilding and textiles, not all firm sizes are known and for the car industry, the concept ofacontinuous distribution is inappropriate for there are only 6 firms. See the note at the end of the Appendix for the methods used to compute $\hat{\mu}_{sj}$ and $\hat{\sigma}^2_{sj}$ in these cases.

from the straight lines drawn through the points. $\hat{\sigma}_{sj}$ is calculated as the inverse of the slope and $\hat{\mu}_{sj}$ as the value of log S at which the normit = 0[1]. Maximum likelihood or minimum normit χ^2 could have been used to calculate these parameters, but there seems little point in such sophistication, given the low number of observations for most industries. (See section 2 of chapter 6 and section 1 of chapter 7.) These figures provide a rough first test of lognormality: the straight line appears to provide a good fit in most cases, but clearly, the brick industry provides an exception.

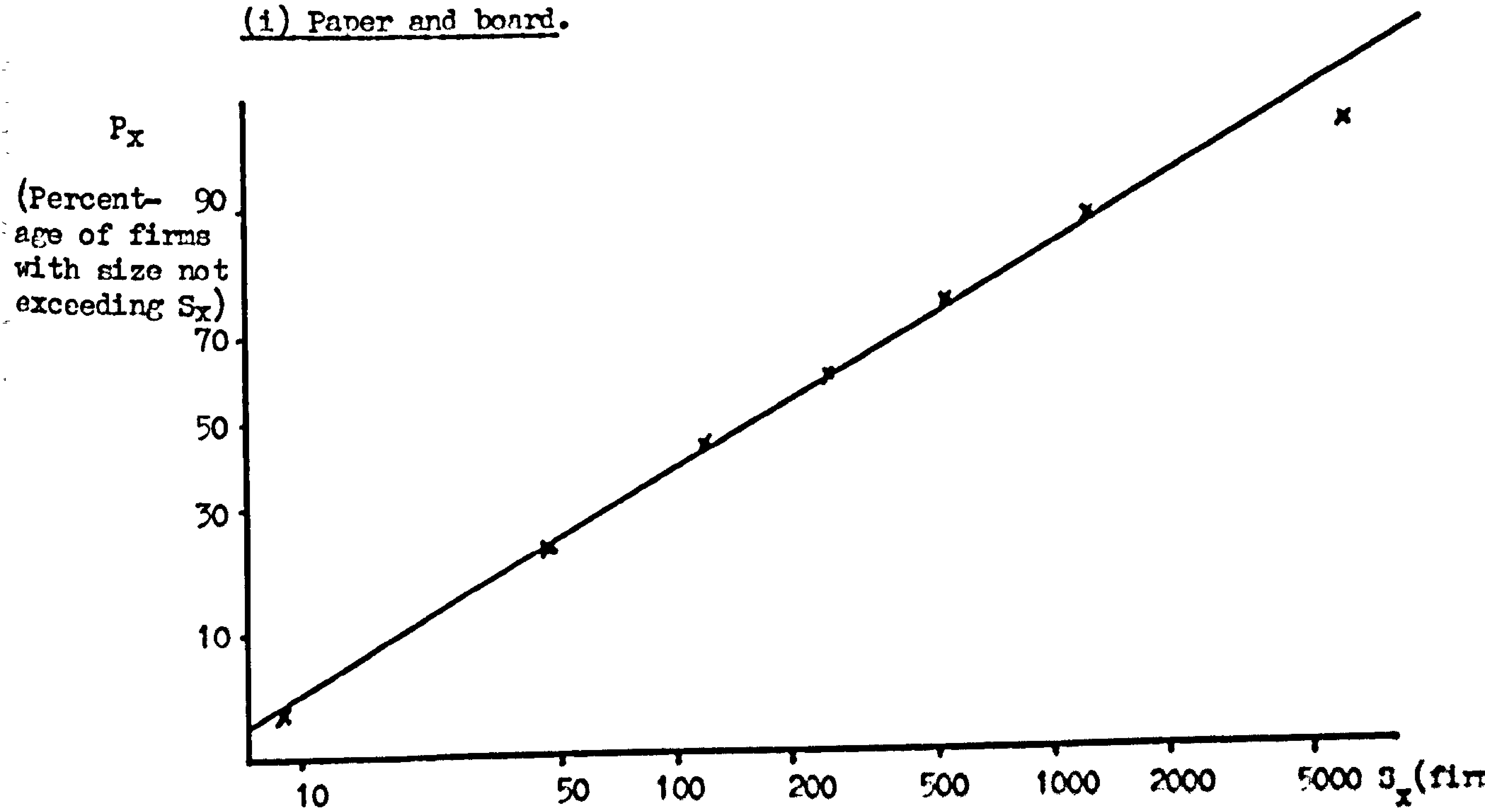
As alternative tests, χ^2 and the Kolmogorov - Smirnoff test statistic have been computed for each industry. The former is, of course, the standard test of normality and is based on the observed and expected numbers of firms in each size class. As can be seen from table A5.2.1. the hypothesis of a lognormal size distribution can only be rejected at the 5% level for the brick industry, although local newspapers, too, record a high value for χ^2.

Whilst this is an extremely encouraging result, it must be admitted that the χ^2 test is fairly imprecise.[2] Therefore, a second test is probably desirable; there are, in fact, a wide variety of tests which may be applied,[3] most of which have been mentioned already. In the event, the so-called Kolmogorov - Smirnoff approach is used. Lilliefors[4] claims

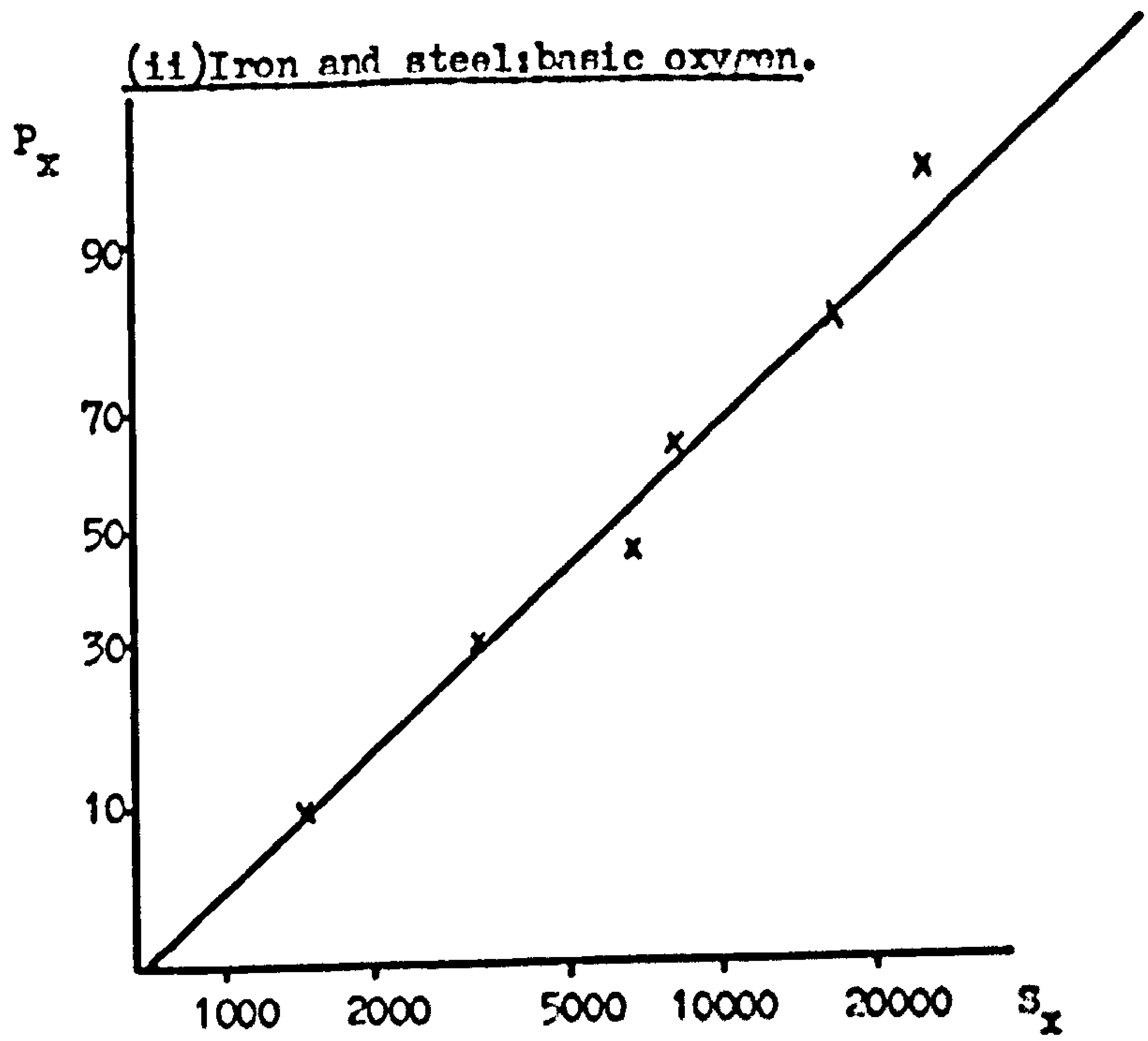
1. If S is $\Lambda(\mu_{sj}, \sigma^2_{sj})$ then the proportion of firms having size less than S_x is given by $P_x = \Lambda(S_x | \mu_{sj}, \sigma^2_{sj}) = N(\log S_x | \mu_{sj}, \sigma^2_{sj})$
Thus $Z_x = -(\mu/\sigma)_{sj} + (1/\sigma_{sj}) \log S_x$ where z_x is the normit of P_x
2. See Quandt's (op.cit.) criticisms.
3. Silbermann's concentration ratio approach, Quandt's statistic , symmetry of Lorenz curves, Fisher's test of normality etc.
4. H.W. Lilliefors, 'On the Kolmogorov - Smirnoff test for normality with mean and variance unknown,' Journal of the American Statistical Association 1967. See also F.J. Massey, "The Kolmogorov-Smirnoff test for goodness of fit, "Journal of the American Statistical Association, 1951.

Figure A.5.2.1: The observed firm size distributions.

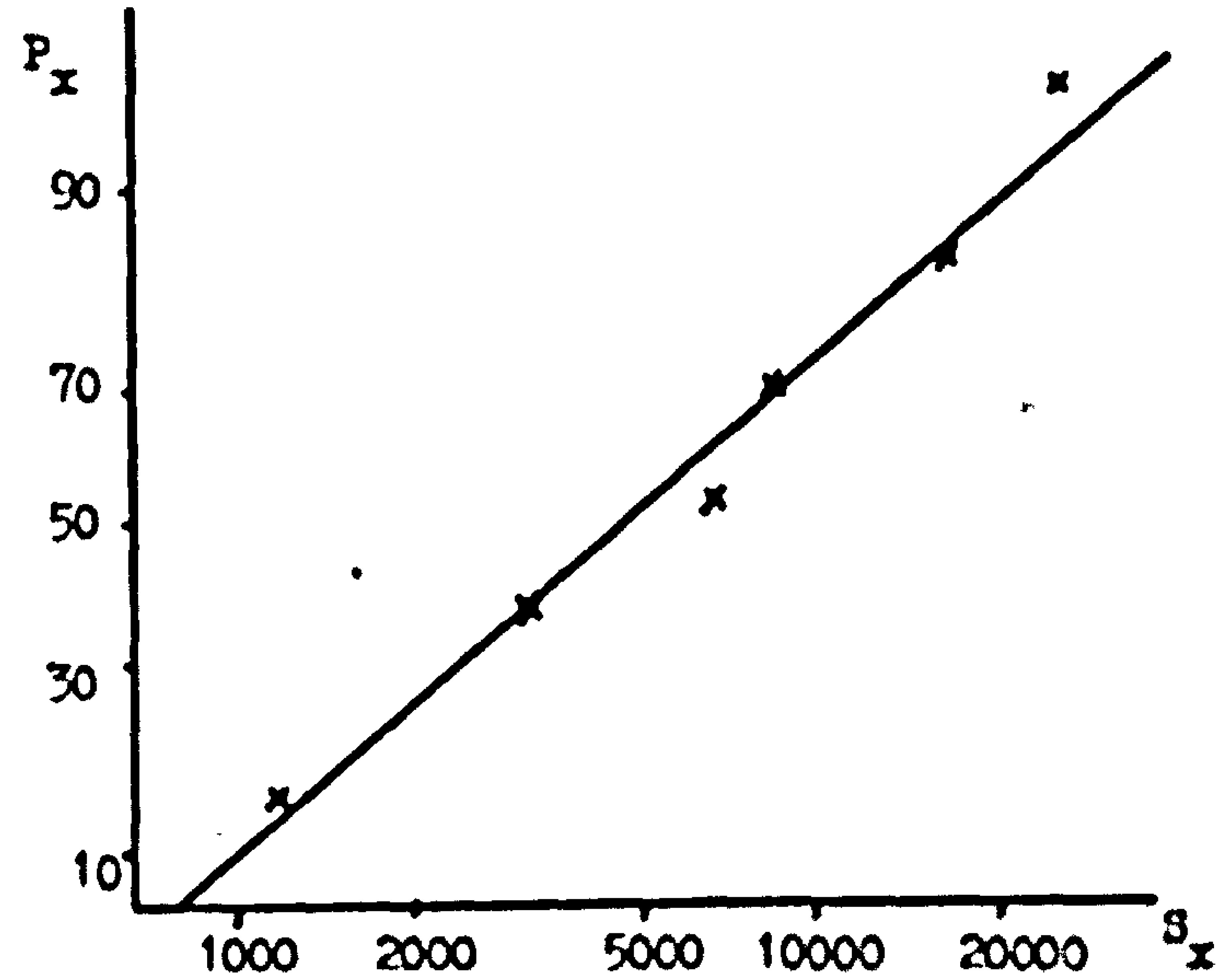
(i) Paper and board.



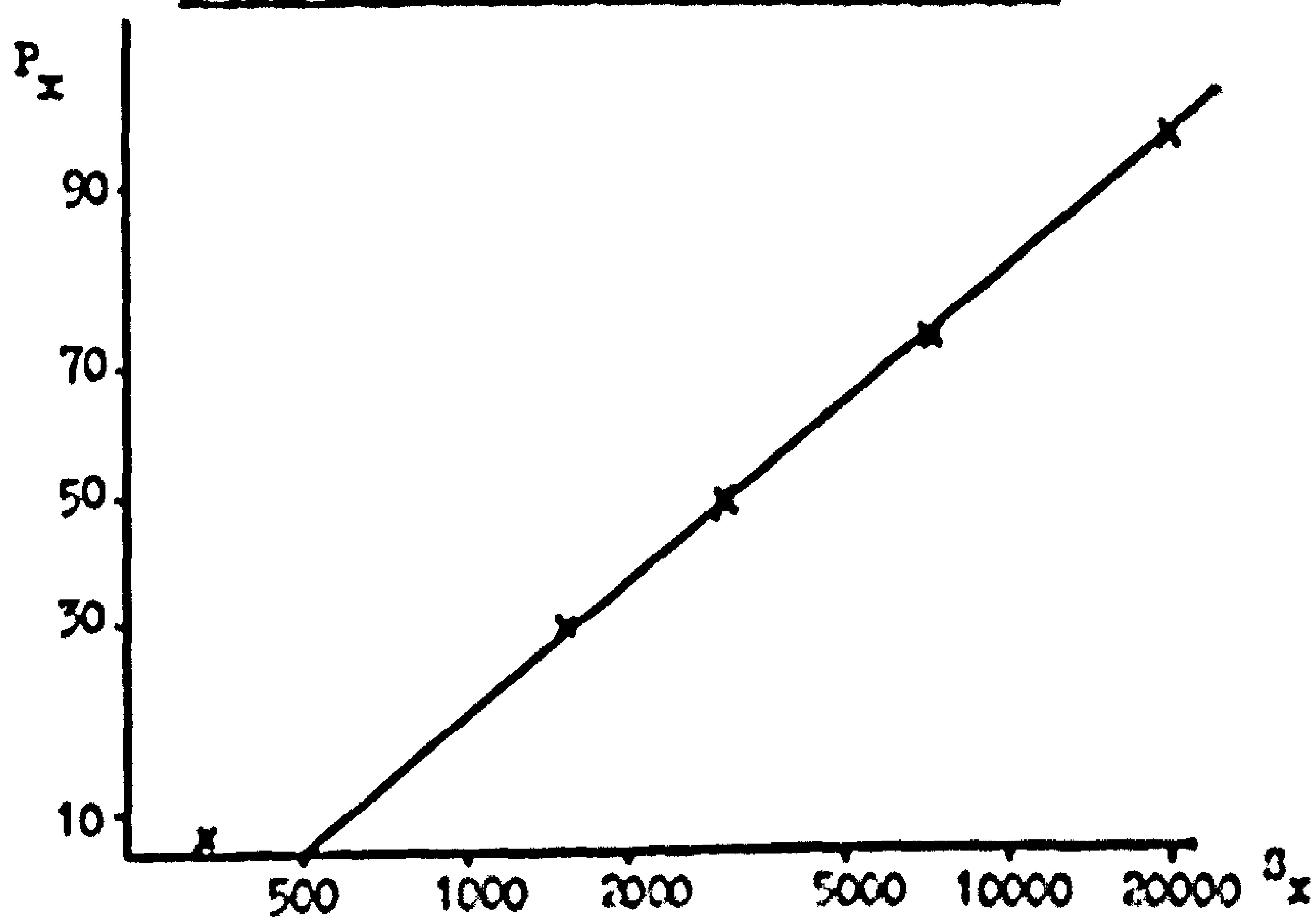
(ii)Iron and steel:basic oxygen.



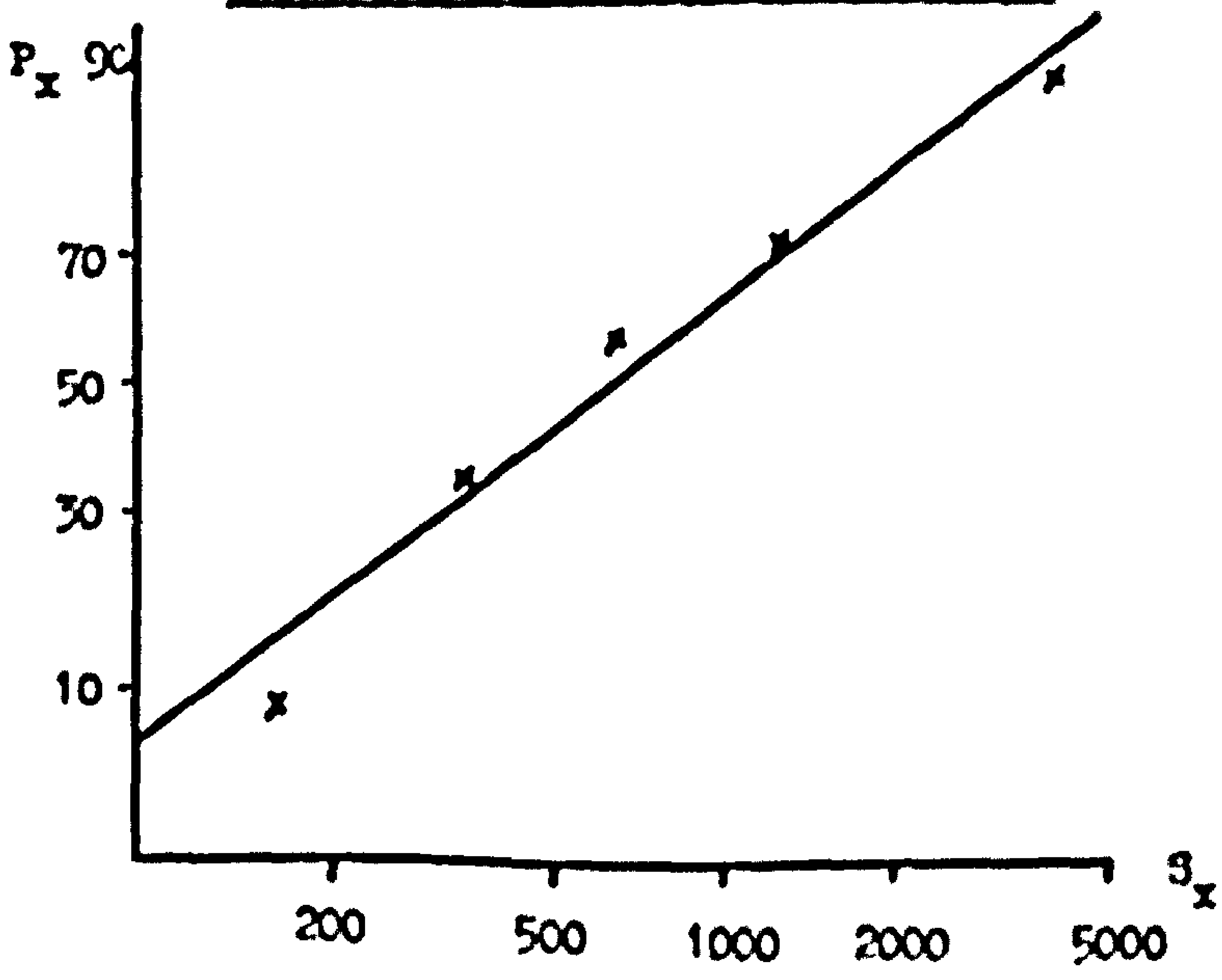
(iii) Iron and steel: continuous casting.



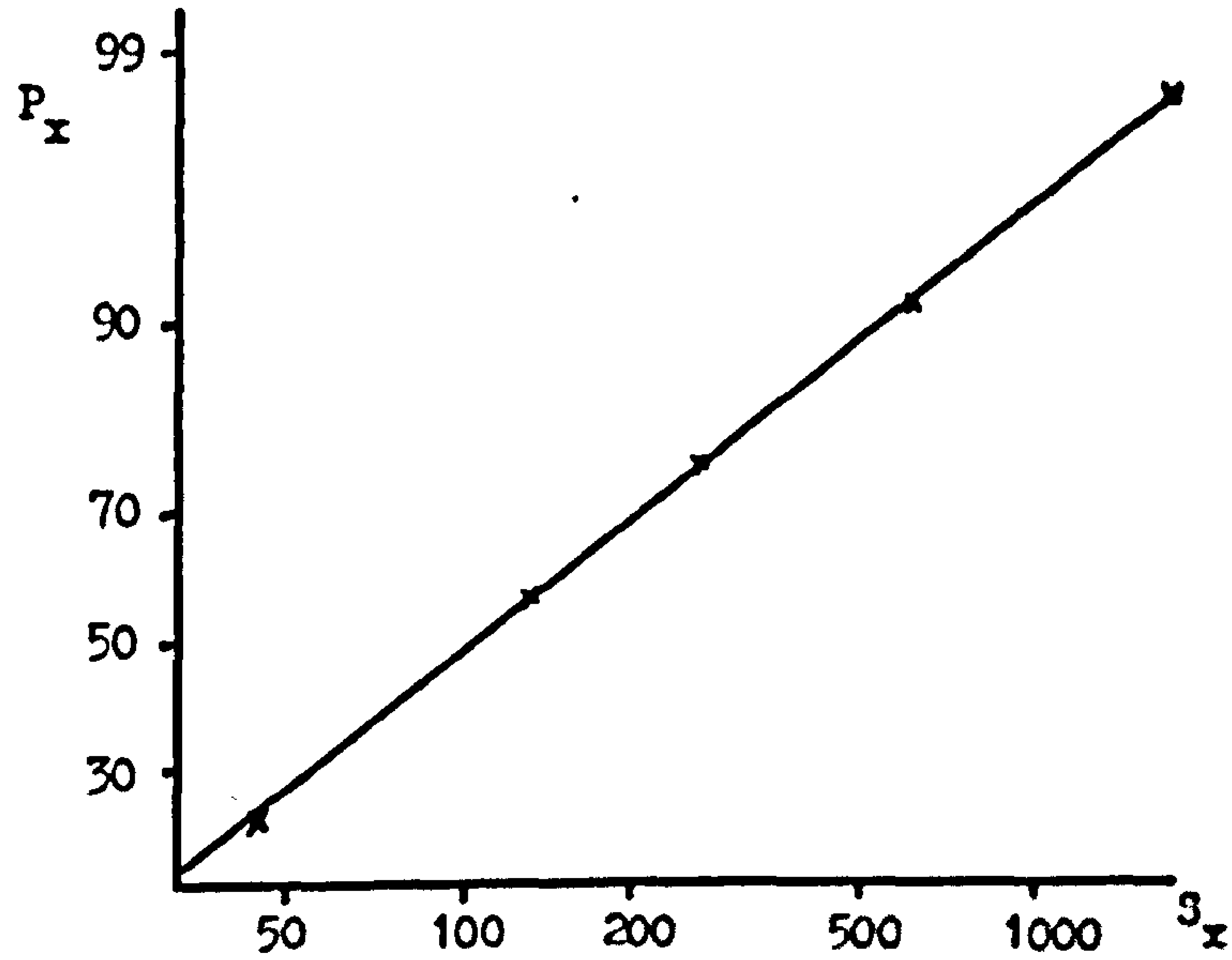
(iv) Iron and steel: vacuum degassing.



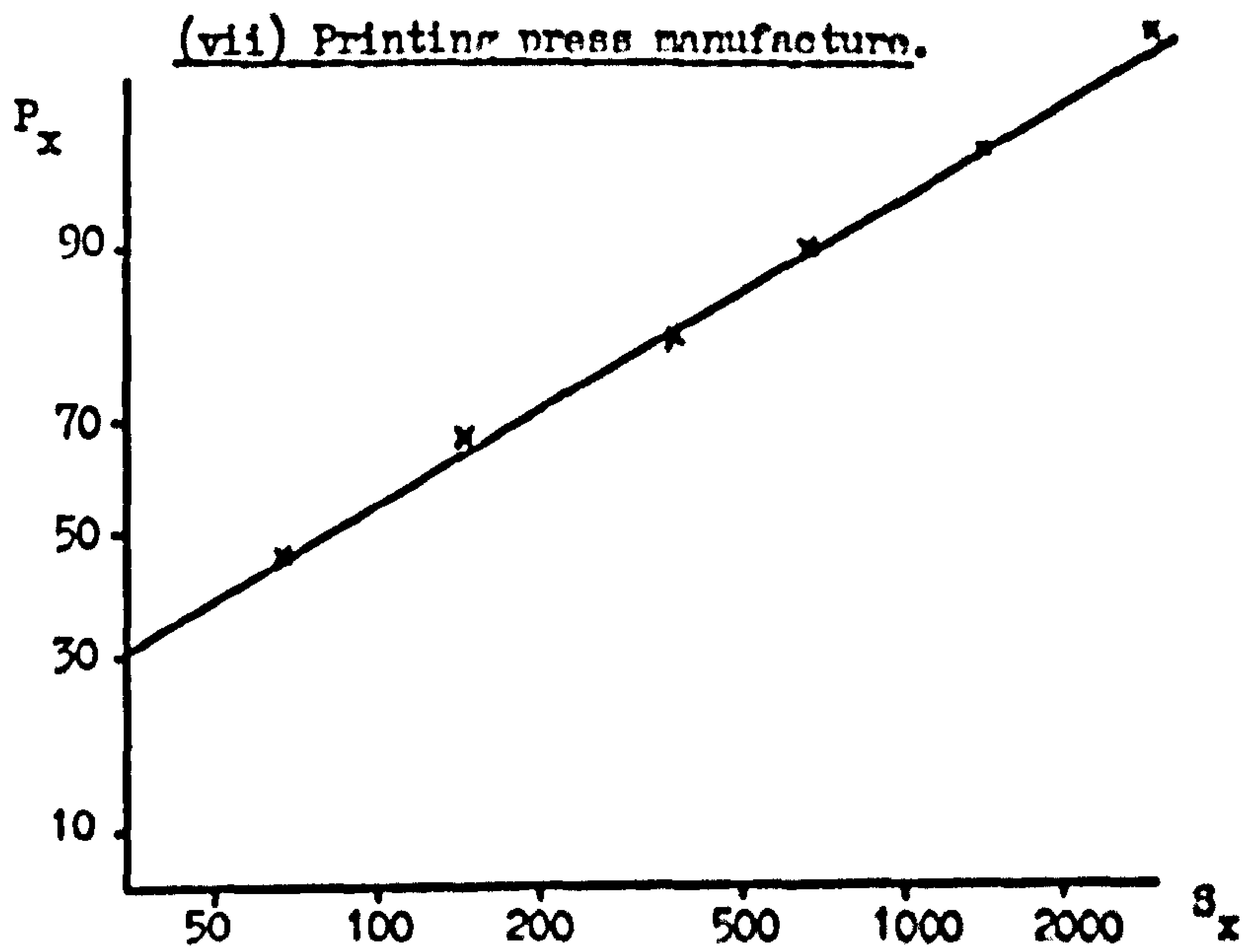
(v) Iron and steel: Vacuum melting.



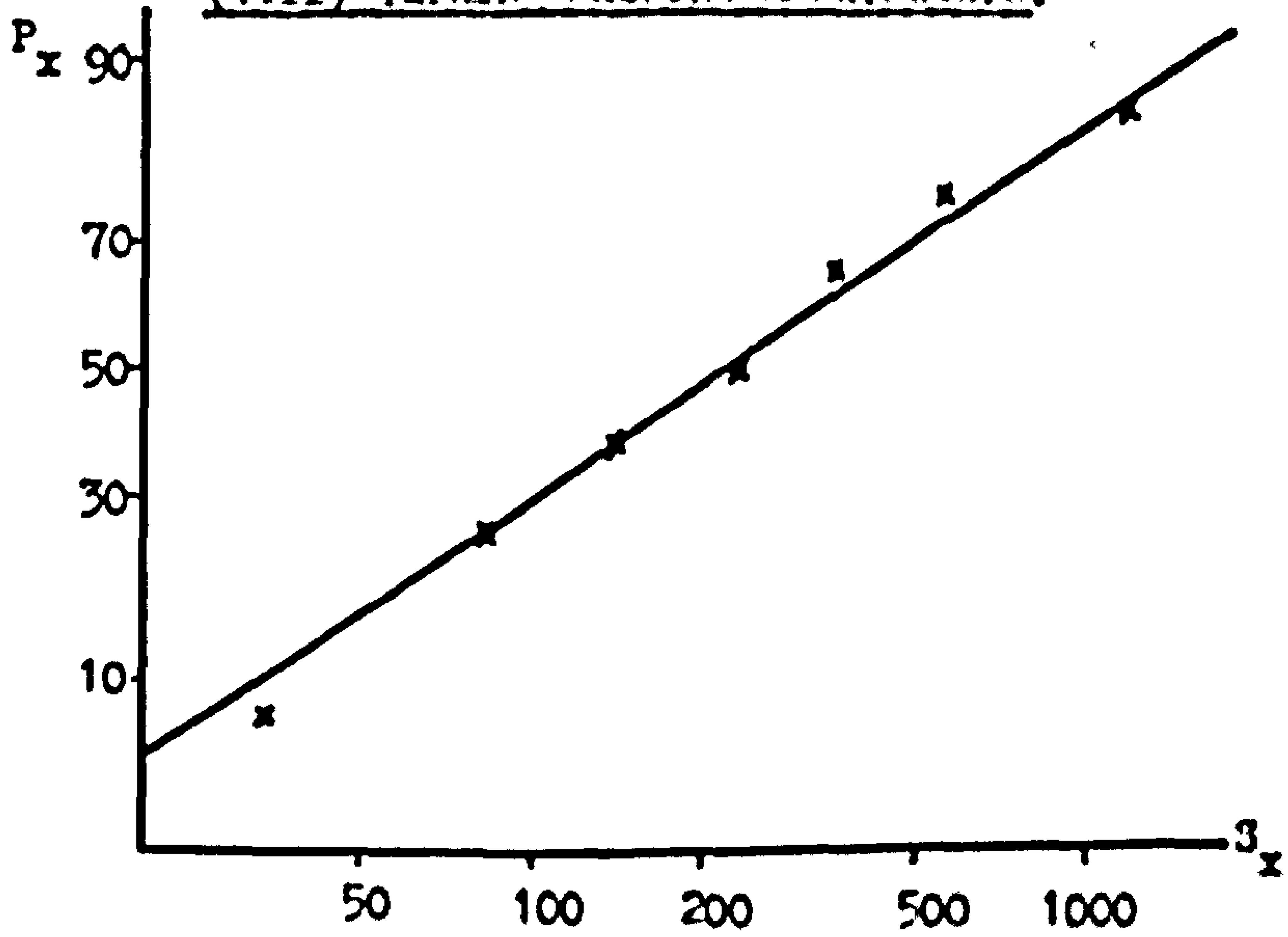
(vi) Lancashire weaving.



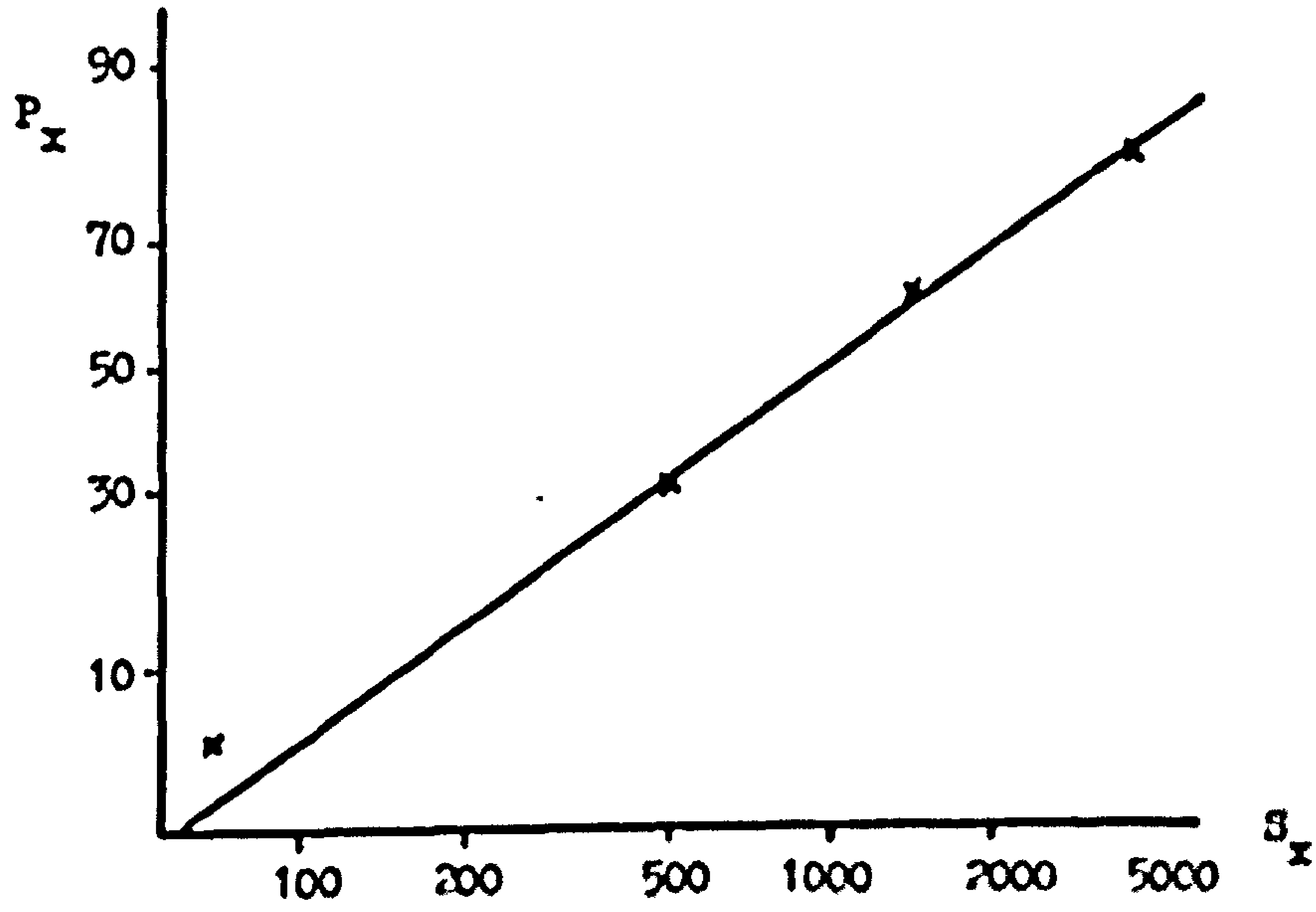
(vii) Printing press manufacture.



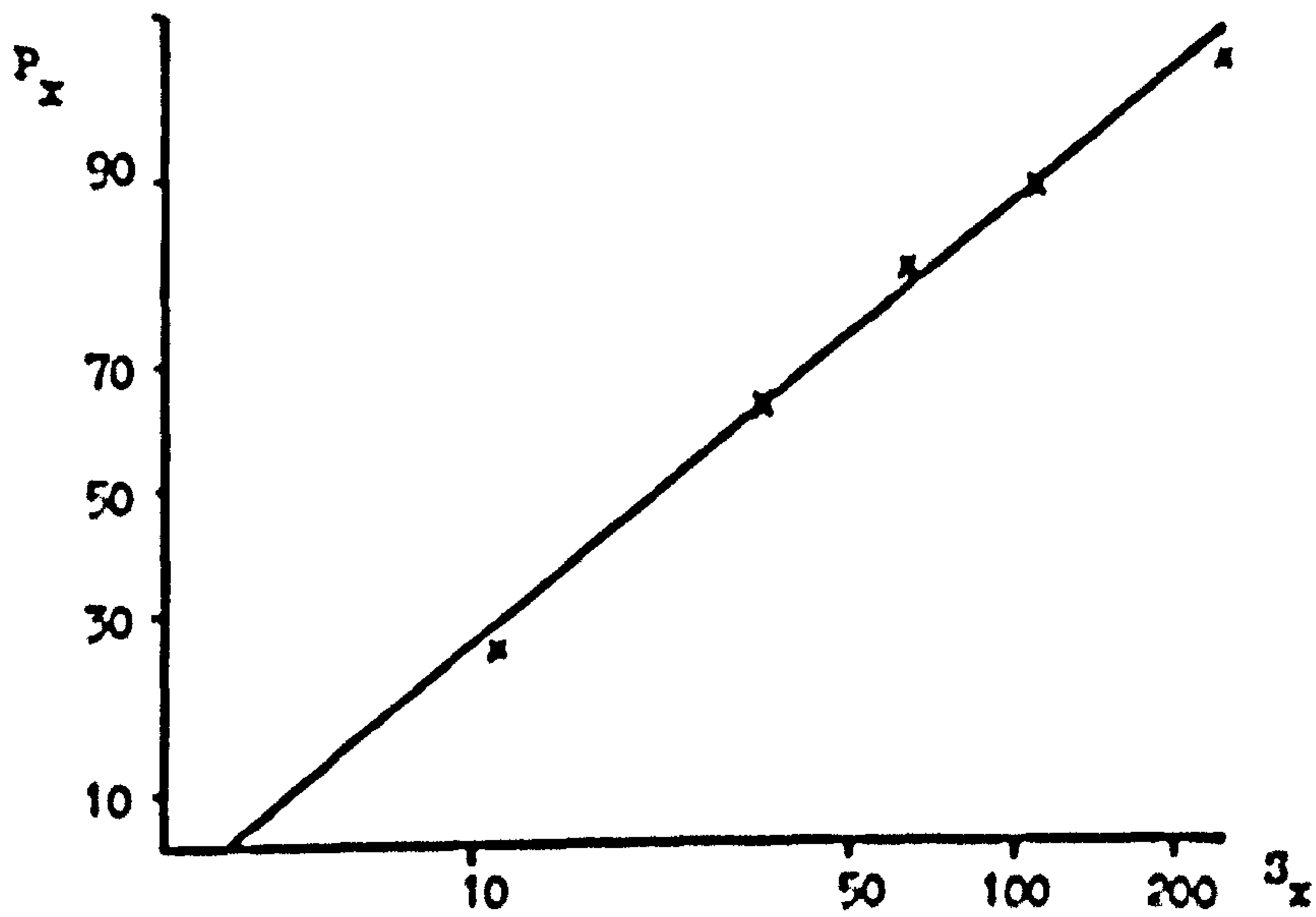
(viii) Turning machine manufacture.



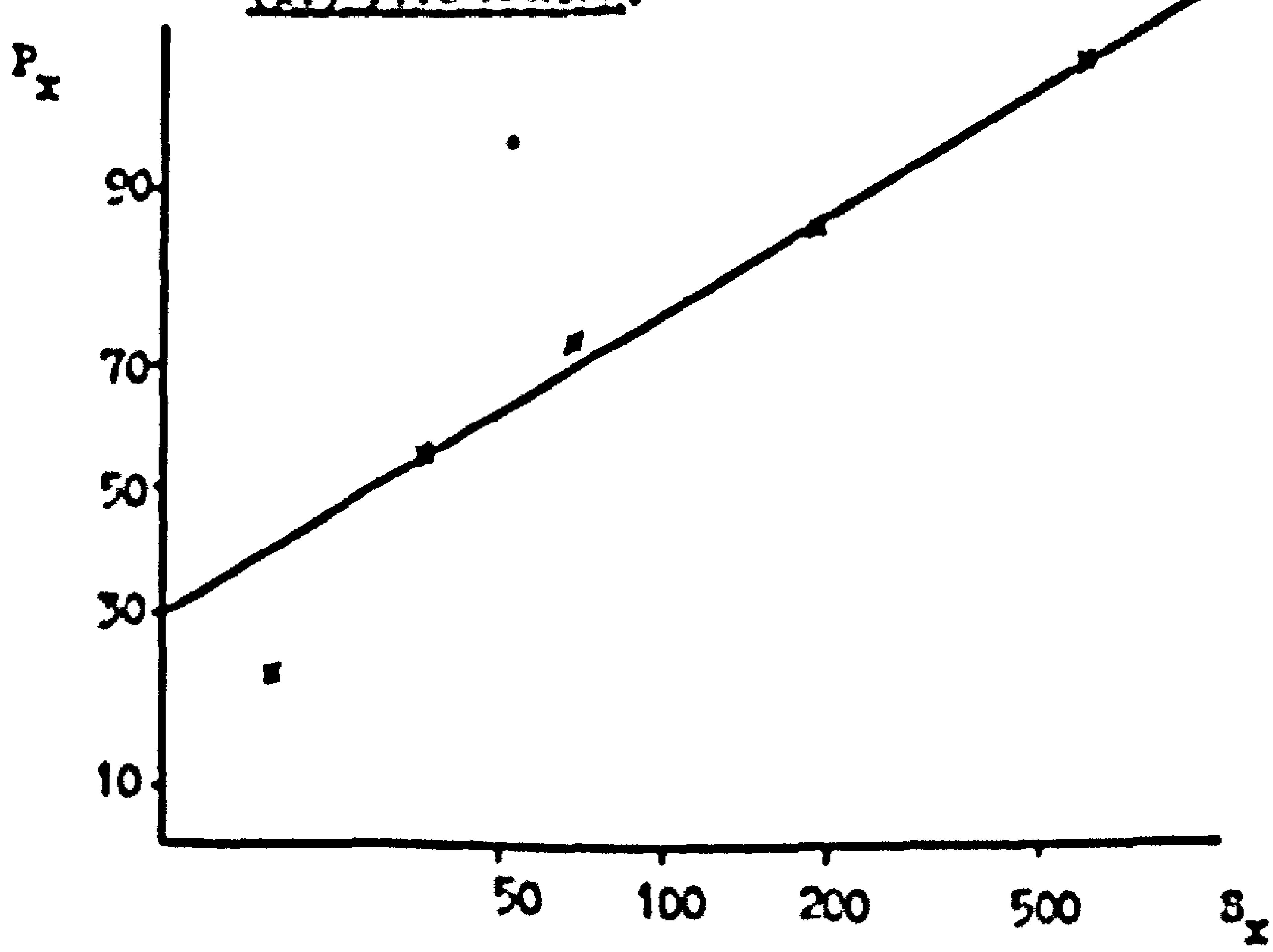
(ix) Turbine manufacture.



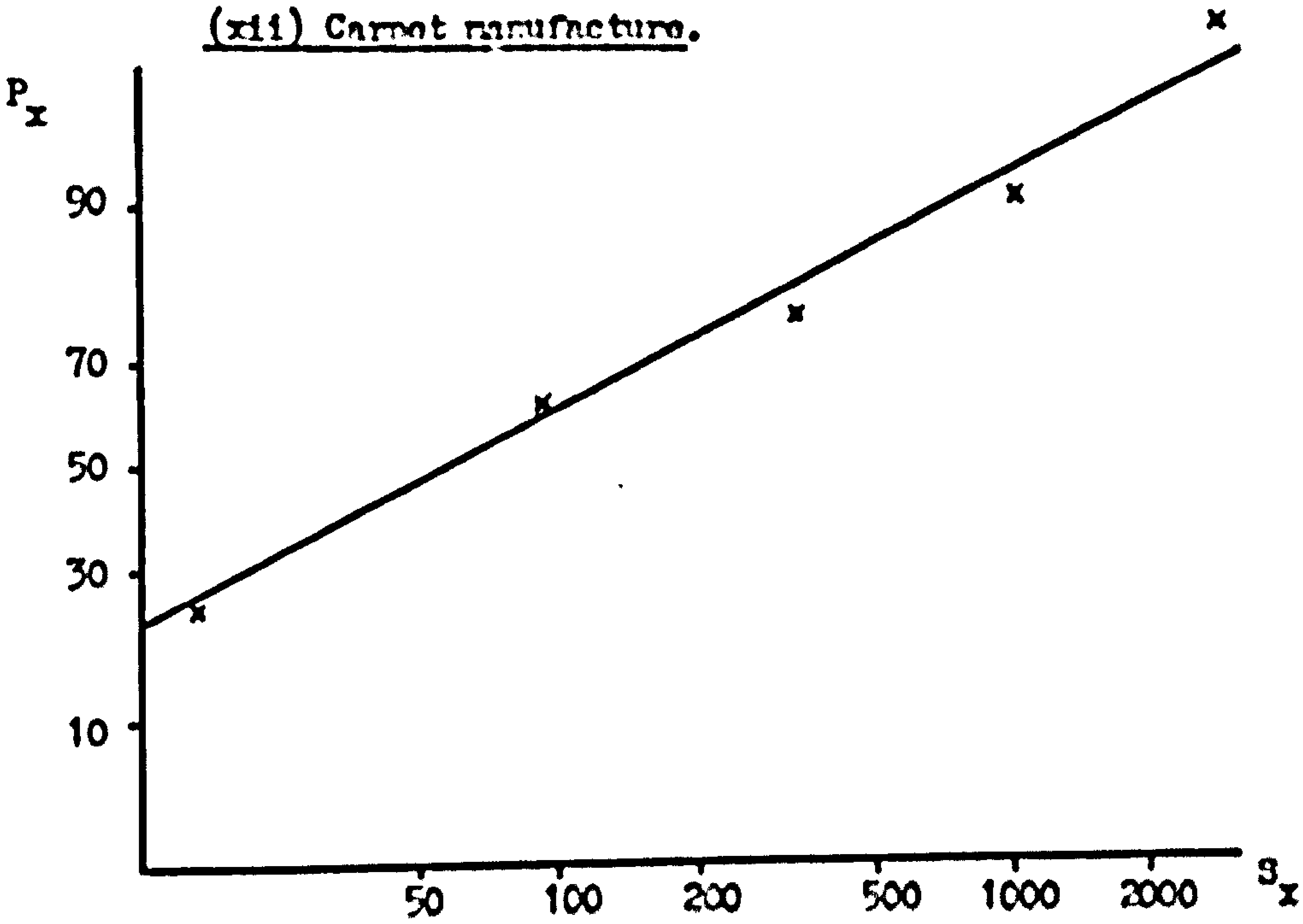
(x) Malting.



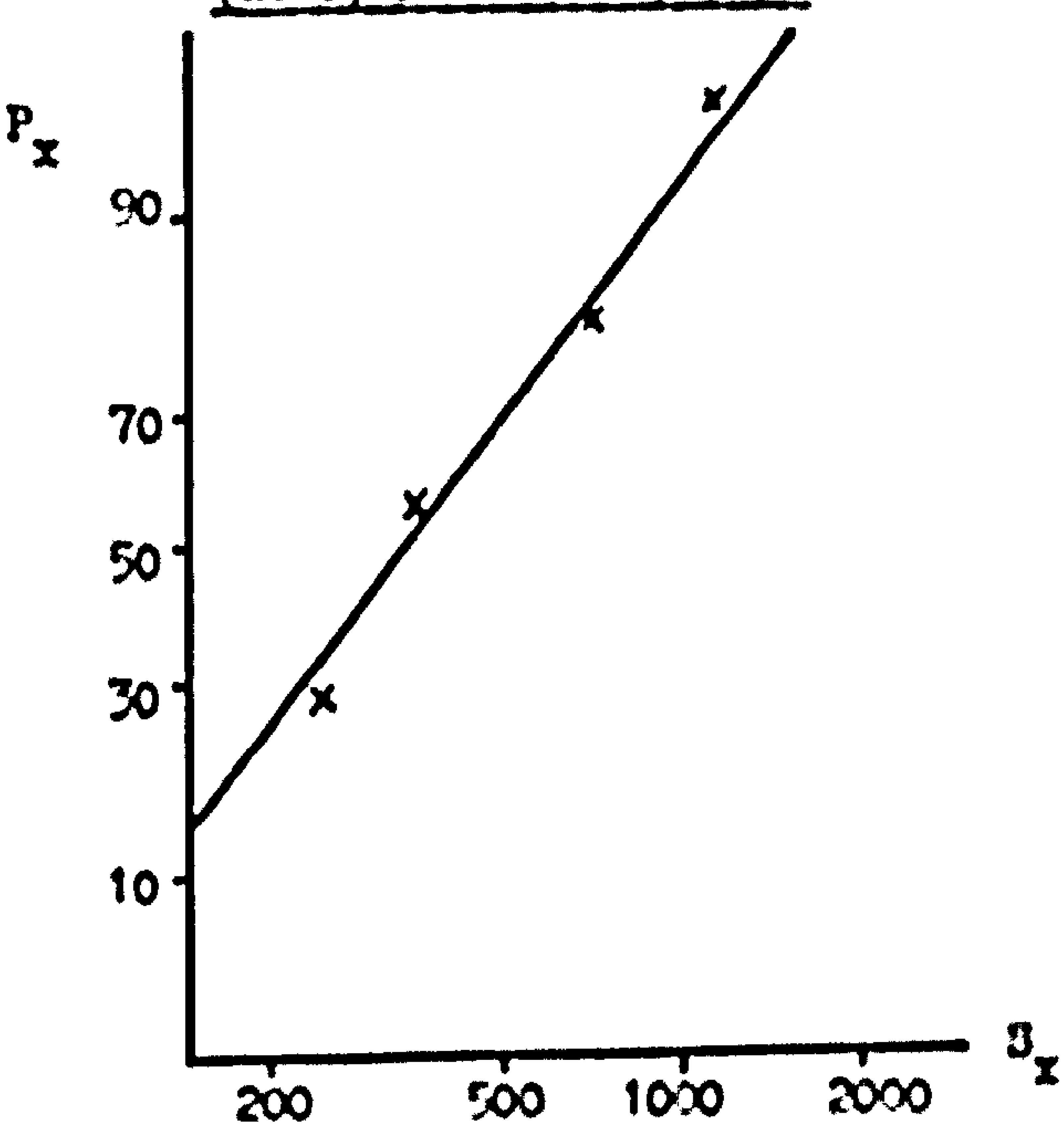
(xi) Brickmaking.



(xii) Carpet manufacture.



(xiii) Local newspapers.



that it may be used with smaller samples where the validity of χ^2 could be questionable, and that for any sample size it will usually be a more powerful test. Basically, it entails a comparison of the observed and predicted cumulative frequencies at each size and then testing to see whether any of the recorded deviations exceed a critical value which would be compatible with normality (for log size), given the total number of firms and confidence level required. More formally, any observed cumulative distribution is rejected as not being normal with x% confidence if $d > d_{cx}$

where $d = \max | O_i - E_i |$

and d_{cx} is the critical value for x% confidence (tables provided in Lilliefors)

O_i is the observed number of firms having size less than S_i

E_i is the predicted " " " " " " S_i

The calculated d_j and d_{cj} for the 5% significance level are presented in table A5.2.1: again, only the brick industry appears to have a size distribution which can be rejected as not lognormal. Indeed, of the other 12 industries, only local newspapers records a value for d which approaches $d_{c,.05}$.

On the basis of these three tests, it seems that the lognormal size distribution assumed in the model is indeed an acceptable approximation. Even for the brick industry, it is suspected that the rejection of the lognormal may result from inaccuracies of data collection. As can be seen from figure A5.2.1.(xi), the straight line only deviates widely from the observed pattern for the smallest size range - it overpredicts badly. This is an extremely surprising result : in virtually all empirical size distributions for which the lognormal has been found to be inappropriate in the past, it has been the upper tail where explanation has been poorest. In this case, it seems quite probable that the data sources used (trade directories) failed to record all small producers of clay bricks, or at least designated them as producers of other kinds of bricks: this would mean that not all small firms are included in the distribution shown for the brick industry,

Table A5.2.1. The estimated parameters and two tests of lognormality.

Industry	$\hat{\mu}_{sj}$	$\hat{\sigma}^2_{sj}$	χ^2_j	K-S Statistics: Observed value (d_j)	K-S Statistics: Critical value (d_{cj})
Paper and Board	5.15	2.941	4.733	2.25	10.3
Local Newspapers	5.82	.603	10.499	6.67	7.8
Malting	3.10	1.531	3.522	2.52	7.7
Lancashire textiles	4.67	1.605	4.964	3.00	24.8
Steel(BOP)	8.68	1.145	6.272	1.86	4.6
Steel (CC)	8.45	1.392	3.504	1.90	4.1
Steel (VD)	8.075	1.501	2.865	1.75	6.4
Steel (VH)	6.5	1.742	5.240	1.83	4.9
Printing press mfg.	5.23	3.063	6.321	3.43	6.7
Turning machine mfg.	5.45	2.236	2.056	2.33	7.4
Turbine mfg.	6.90	1.960	.570	0.56	4.6
Non-fletton Bricks	3.52	2.756	40.634[1]	20.96[2]	12.0
Carpets	4.07	3.842	5.809	4.50	12.3

1. Lognormality rejected at the 5% level.

2. Lognormality rejected at the 5% level.

N.B. Insufficient data is available to test lognormality for Shipbuilding, Car manufacture and Textiles.

resulting in a downward bias in the first point of the figure.

For the car, shipbuilding and textile industries, it has not been possible to test the lognormality assumption; but, nevertheless, estimates of μ_{sj} and σ^2_{sj} are required for these industries by the analysis in the main text. For the car industry, these are computed, simply, as the mean and variance of log size for the six firms. This method generates estimates of μ_s and σ^2_s, for 1967, of 10.16 and .3645 respectively. If size is lognormally distributed, then this would be an acceptable alternative method for deriving the parameters. As a rough check of lognormality, the following test is possible. From the lognormal, the arithmetic mean size is given by $e^{\mu_s+\frac{1}{2}\sigma^2_s}$. Using the above estimates, a prediction of the arithmetic mean size is $e^{10.342}$ which compares with an actual arithmetic mean in 1967 of $e^{10.223}$; thus the observed size distribution appears to exhibit a positive skewness quite close to the lognormal distribution: $\Lambda(10.16, .3645)$.

For shipbuilding, data is not available for the size of each firm. However, it is known that in 1965 (according to the Geddes report)[1] the largest 27 firms accounted for 75% of industry employment. As there were 62 firms in the industry in that year, assuming size to be lognormally distributed, if S^* was the size of the 28th largest firm, then

$$\frac{62-27}{62} = \Lambda(S^*|\mu,\sigma^2) = N\left(\frac{\log S^*-\mu}{\sigma}\,\Big|\,0,1\right) \qquad (A5.2.1)$$

Therefore,

$$\frac{\log S^*-\mu}{\sigma} = .162 \text{ (the normit of } \tfrac{62-27}{62}\text{)} \qquad (A5.2.2)$$

Further, as the first moment distribution, $\Lambda(\mu+\sigma^2,\sigma^2)$, describes the share of total employment accounted for by each firm size

$$.25 = \Lambda(S^*|\mu+\sigma^2,\sigma^2) = N\left(\frac{\log S^*-\mu}{\sigma}-\sigma\,\Big|\,0,1\right) \qquad (A5.2.3)$$

Therefore $\frac{\log S^*-\mu}{\sigma} - \sigma = -.675$ (the normit of .25) $\qquad$ (A5.2.4)

Subtracting (A5.2.4) from (A5.2.2), $\sigma = .837$ and $\sigma^2 = .7006$.

As the arithmetic mean size in 1965 was 1048, again assuming lognormality

$$\mu = \log 1049 - {}^1/_2\sigma^2 = 6.9546$$

1. op.cit.

The position is exactly the same for weaving : individual firm sizes are not known, but it is known that in 1969, the top 12 firms accounted for half of employment (Textile Council Report)[1]. As there were 279 firms in the industry with a mean size of 168 employees in that year,

$$\frac{267}{279} = N\left(\frac{\log S^* - \mu}{\sigma} \,\middle|\, 0,1\right) \quad \text{where } S^* \text{ is the size of} \qquad (A5.2.5)$$

the 13th largest firm.

Therefore:

$$\left(\frac{\log S^* - \mu}{\sigma}\right) = 1.7 \text{ (the normit of } \frac{267}{279}\text{)} \qquad (A5.2.6)$$

From the first moment distribution, $.50 = N\left(\frac{\log S^* - \mu}{\sigma} - \sigma \,\middle|\, 0,1\right)$ (A5.2.7)

therefore $\frac{\log S^* - \mu}{\sigma} - \sigma = 0$ (A5.2.8)

Thus, $\sigma = 1.7$, $\sigma^2 = 2.89$, and $\mu = \log 168 - 1.445 = 3.679$

To summarise then, for each of these three industries it is necessary to assume lognormality, in which case the estimated parameters are as in table A5.2.2.

Table A5.2.2. Estimated parameter using other methods.

Industry	$\hat{\mu}_{sj}$	$\hat{\sigma}^2_{sj}$
Cars	10.16	.3645
Shipbuilding	6.9546	.7006
Weaving	3.67896	2.89

3. Constancy of the parameters.

Finally, there remains the problem of whether it is acceptable to assume, for all industries, that μ_{sj} and σ^2_{sj} are invariant with respect to time. Unfortunately, for most firms in the sample industries, data on size is available only for one point in time (usually the year of data collection.) Similarly where published size distributions have been used, these too are not available for more than one year. Thus, for most industries, $\hat{\mu}_{sj}$ and $\hat{\sigma}^2_{sj}$ are only available for one point in time.[2]

It seems unlikely that σ^2_{sj} will change by very much over the diffusion

1. op.cit.

2. The distribution published in Census reports are at a too aggregated level to be useful here.

period, especially as firms entering or leaving during the period have been excluded in all cases. Indeed, for the only sample industry for which distributions are recorded at the start and end of the period (Lancashire textiles), σ_s^2 changed by less than 5% over 20 years.[1] On the other hand, a constant μ_s seems less likely. Fortunately, the implications are less serious than for non-constant σ_s^2 : in the time series econometrics, this may be handled fairly easily (see Appendices at the end of chapters 5 and 6), and in the cross-industry empirics it is unlikely that the relative sizes across industries of μ_{sj} will change radically over the period. This problem can be seen in perspective by the fact that the average growth in employment over the diffusion period for the sample industries was less than 1% per annum; with constant σ_{sj}^2, this implies an average change in $e^{\mu_{sj}}$ over a 10 year period of only about 10%.

1. Computed from data presented in C. Miles, 'Lancashire textiles: A case study of industrial change.' Cambridge University Press, 1968. For some general evidence, see L. Engwall, (op.cit.) particularly p.75.

4. Empirical size distributions.

Table A5.4.1. Data used in figure A5.2.1.

(i) Paper and Board (1967)

Size range[1]	Number of firms.	Mean firm size(employees)[2]
1 - 999 tons	8	9
1000 - 4999 tons	27	47
5000 - 9999 tons	16	122
10000 - 19999 tons	15	246
20000 - 49999 tons	18	531
50000 - 99999 tons	6	1268
≥ 100000 tons	8	6391
Total	98	768.4

1. In tons of paper and board produced.

2. The published data makes no reference to employment. This column has been calculated assuming that all size ranges have average labour-output ratios equal to the labour output ratio for the industry as a whole.

Source: British Paper and Board Manufacturers Association Reference tables.

(ii) Iron and Steel - for use of Basic Oxygen Process (1964)[1]

Employment range	Number of firms	Average employment
1 - 3999	8	2298
4000 - 7999	3	6617
8000 - 9999	4	8263
10000 - 17999	3	16829
≥ 18000	2	25396
Total	20	8628

1. i.e. the year of nationalization.

Sources: Iron and Steel Board Annual Statistics and Trade directories.

(iii) Iron and Steel – for use of continuous casting (1964)[1]

Employment range	Number of firms	Average employment
1 – 1999	7	1223
2000 – 3999	4	3195
4000 – 7999	4	6738
8000 – 9999	4	8263
10000 – 17999	3	16829
≥ 18000	2	25396
Total	24	7609

1. i.e. the year of nationalisation.

Sources: Iron and Steel Board Annual Statistics and Trade directories.

(iv) Iron and Steel – for use of vacuum degassing (1964)[1]

Employment range	Number of firms	Average employment
1 – 999	6	338
1000 – 1999	9	1516
2000 – 4999	7	3025
5000 – 9999	11	7183
≥ 10000	5	20256
Total	30	5714

1. i.e. the year of nationalisation.

Sources: Iron and Steel Board Annual Statistics and trade directories.

(v) Iron and Steel – for use of vacuum melting (1969)

Employment range	Number of firms	Average employment
1 – 199	4	157
200 – 499	7	330
500 – 999	3	633
1000 – 1999	3	1267
≥ 2000	5	4140
Total	22	999

Source: British Iron and Steel Producers Association (BISPA) directory.

(vi) Lancashire weaving (1956)

Size range[1]	Percentage of total firms	Average employment[2]
1 - 200	48	45
201 - 400	19	136
401 - 800	19	272
801 - 2000	11	634
≥ 2000	3	1333

1. Numbers of looms employed.

2. Based on an average loom-labour ratio derived from data provided in Robson op.cit.

Source: J.S. Metcalfe, op.cit., p.172.

(vii) Manufacture of printing presses (1970)

Employment range	Number of firms	Average employment
1 - 49	8	36
50 - 99	9	65
100 - 299	8	144
300 - 499	6	366
500 - 999	5	650
1000 - 1999	3	1370
≥ 2000	4	2700
Total	43	527

Sources: Trade directories, Kompass, N.I.E.S.R. sampling frame.

(viii) Manufacture of turning machines (1970)

Employment range	Number of firms	Average employment
1 - 49	10	33
50 - 99	8	81.5
100 - 199	4	140
200 - 299	9	227.9
300 - 499	7	350
500 - 749	4	537
750 - 1999	1	1179
⩾ 2000	4	4000
Total	53	614

Sources: Trade directories, Kompass and N.I.E.S.R. sampling frame.

(ix) Manufacture of turbines (1970)

Employment range	Number of firms	Average employment
1 - 499	6	172
500 - 999	4	737
1000 - 1999	3	1400
⩾ 2000	7	3650
Total	20	1606

Sources: Trade directories, Kompass and N.I.E.S.R. sampling frame.

(x) Malting (1967)

Employment range[1]	Number of firms	Average employment
1 - 24	28	11
25 - 49	16	34
50 - 99	5	63
100 - 149	3	109
≥ 150	4	256
Total	56	44.8

1. Excluding personnel employed on brewing for firms that brew and malt.

Sources: N.I.E.S.R. questionnaires, Kompass and Brewers' Almanack.

(xi) Non-Fletton bricks (1970)

Employment range	Number of firms	Average employment
1 - 24	58	19
25 - 49	33	37
50 - 99	19	69
100 - 249	16	194
≥ 250	11	620
Total	137	99

Sources: N.I.E.S.R. Questionnaire, Kompass, Trade directories.

(xii) Carpet manufacture (1968)

Employment range	Number of firms	Average employment
1 - 24	67	16
25 - 199	21	95
200 - 499	17	326
500 - 1999	16	1025
≥ 2000	4	2064
Total	125	214

Source: Census of production, [illegible] report No. 419, excluding 16 new entrants in the 25 - 199 range on the basis of information received from T. Scott (op. cit.)

(xiii) Local newspapers.

Employment range	Number of firms	Average employment
1 - 299	24	242
300 - 399	11	323
400 - 699	10	493
700 - 999	8	707
≥ 1000	4	1450
Total	57	462

Source: Evening newspaper advertising bureau, Kompass and Kelly's directory.

These data are plotted in cumulative form in figures (5.2.1) to (5.2.13), in which observations refer to the mid-points in each range and the mean size in that range.

Appendix 6. Problems in measurement of Market Structure.

In chapters 8 and 9, much of the analysis relates to the influence of market structure on the observed structural parameters of the model. This appendix discusses, at some length, the most appropriate empirical measures of market structure and concludes that, for any industry, the parameters of the firm size distribution provide a three-dimensional reflection of structure, which has certain advantages over the more traditional measures.

Section one repeats briefly the arguments for including the level of competition (CC_j) as an influence on the diffusion process. Section two translates these arguments into a series of characteristics which the chosen empirical measure should reflect. Section three surveys the conventional measures of industrial concentration. Section 4 discusses their relative advantages and limitations, on the assumption that firm size is lognormally distributed. Section 5 derives a generalised measure of concentration, the U index, which appears to be more flexible than any of the conventional measures. Section 6 discusses briefly, the alternative measures of barriers to entry and product differentiation.

1. The role of industrial structure (CC_j) in the diffusion process.

From section 6 of chapter 5, which itself summarises the arguments of chapter 4, CC_j is postulated as a determinant of the five structural parameters: α_j, β_j, 6_j, Ψ_j and Ω_j in industry j's diffusion process.

More specifically, it has been argued that market structure may influence:[1]

1) the competitive pressures on non-adopters to adopt.

2) the incentives for non-adopters to adopt.

3) adopters' willingness to pass on information about the new innovation to non-adopters.

1. Often, one cannot be certain as to the direction of these influences however.

4) the attitudes held by all firms to new ideas and innovations.

5) the homogeneity of products offered by an industry.

In terms of the model, pressures will influence initial attitudes, propensity to search and the variability in attitudes (as reflected in α_j, Ψ_j and Ω_j and σ_j respectively.) Incentives should be represented by the typical profitability of the innovation, which will influence also propensity to search, returns from search, (as reflected in Ψ_j, Ω_j and α_j). Willingness to pass on information will partially determine the returns from search (as reflected in Ψ_j and Ω_j). General attitudes (which will also partly depend on pressures) will influence investment yardsticks, propensity to search and perhaps the tendency for larger firms to adopt earlier[1] (as reflected in α_j, Ψ_j, Ω_j and β_j.) Finally, the greater the homogeneity of the industry's product, the less variability there is likely to be across firms in the profitability of the new innovation (as reflected in σ_j.)

2. Three dimensions of market structure.

These arguments suggest that at least three dimensions of market structure may be relevant. It is conventional to argue that in industries in which a large share of the market is concentrated[2] in a few firms there is more chance of implicit (or even explicit) collusion, with consequent lack of competitive pricing and perhaps a tendency towards maintaining the status quo. Thus, factors 1, 3, 4 and 5 from the list above may be influenced by the level of industrial concentration. The two constituent parts of concentration are, of course, the number of firms in the industry and the inequalities between their sizes: the more firms there are in the industry, the less chance there is of collusion, so long as a few firms do

1. See the earlier reference (section 3 of chapter 2) to Adams and Dirlam's finding in this area.

2. See Scherer, op.cit., p.50, for instance. The dimensions to be discussed here follow the scheme suggested by J. Bain, 'Barriers to new competition', Harvard University Press, Cambridge, Mass. 1956.

not dominate the market. Conceptually, in an industry of, say, 100 firms, collusion is more probable if the 5 largest have a market share of 80%, than if their market share is only 10%. On the other hand, an industry with the top 10 of 200 firms accounting for 80% of the market should be more competitive, in this sense, than the above example of the top 5 of 100 firms accounting for 80%.[1] Secondly, the extent of barriers to entry may also determine the degree of management slack and the size of the mark-up used in pricing and, as such, may influence all five factors in the above list. The usual barriers considered are scale economies, production differentiation, advertising expenditures and minimum capital requirements. Thirdly the extent of product differentiation is an important dimension of structure in its own right. Where firms have successfully differentiated their product, both the own price and cross-price elasticities will be low, permitting large mark-ups. Again, this will influence all five factors in the list but particularly, perhaps, pressures on non-adopters to adopt and homogeneity of products. There are always serious problems in measuring this dimension.

3. Alternative measures of concentration.

In past research a wide range of alternative statistical measures of concentration have been used, in this section, five of the more common ones are surveyed briefly and contrasted.[2]

(i) The concentration ratio (until recently, by far the most popular) measures the share of total industry employment[3] accounted for by the 5 largest firms in the industry.[4] Two useful attributes of the concentration

1. The 5% proportional concentration ratio is 80% in both cases.

2. It has been claimed that all measures of concentration are so highly correlated that it matters very little which one is chosen for empirical purposes.(e.g.Scherer, op.cit., p.52). Whilst the argument to be presented shows that these measures do have much in common, there are also significant differences which may be crucial in some cases.

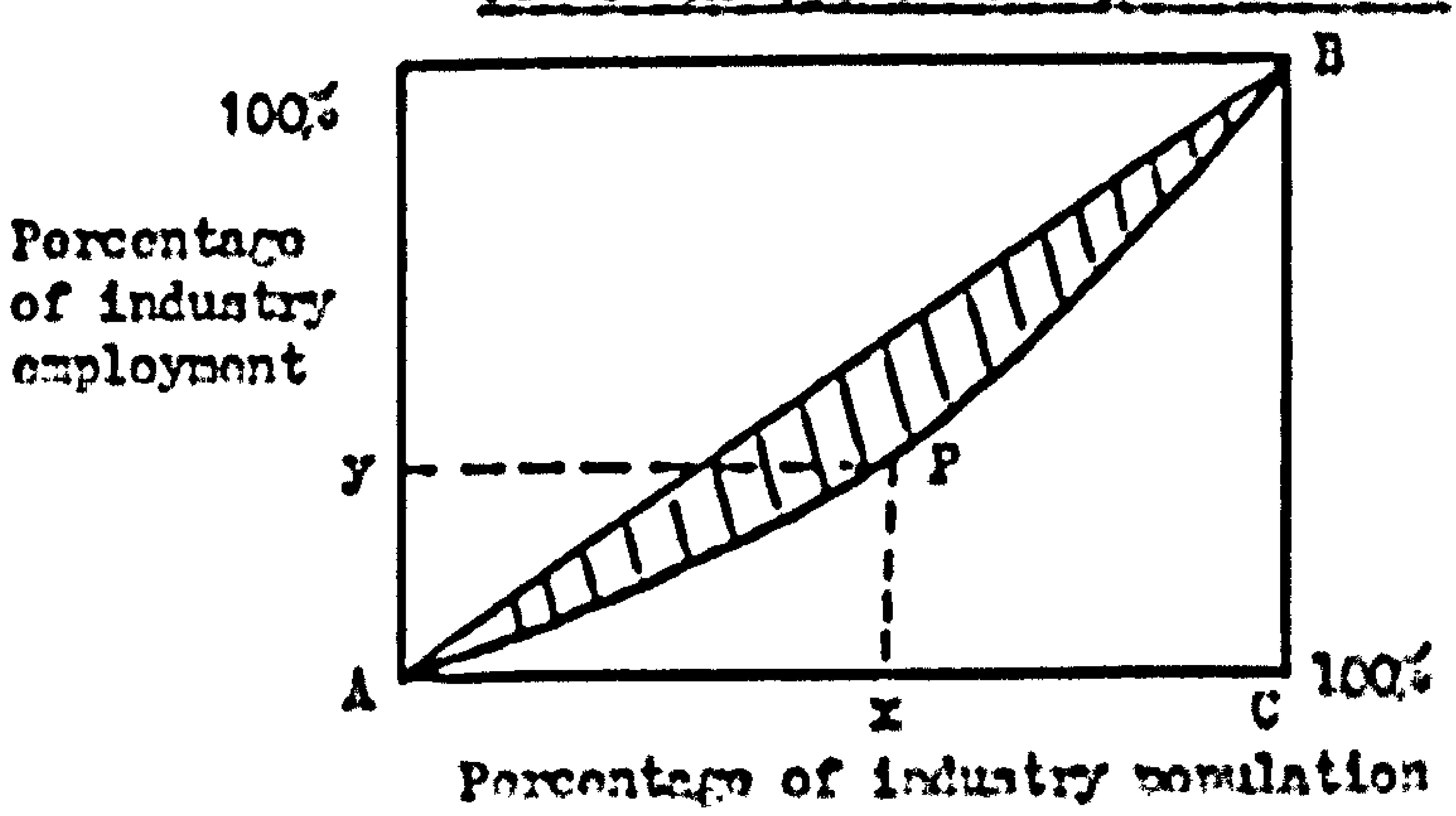
3. In this study, firm size is measured by employment, merely because data is more readily available than for output, sales, value added, assets etc. Needless to say, each of these units has its drawbacks - employment probably tends to understate the size of the largest firms(assuming they are more capital intensive.)

4. Although often 3,4,8 or 20 firm concentration ratios are used.

ratio (CR_5) are that it focusses attention on the all-important upper tail of the size distribution and that it is sensitive to both the total number of firms and the inequality in their sizes. A major limitation is that it relates to only one point on the cumulative concentration curve: there is nothing God-given about the number five and, in many industries, it might be more meaningful to examine the employment shares of individual firms within and outside the top 5. For example, an industry in which the largest firm employs 30% of the total work force and 70 others each employ 1% has a similar CR_5 to one in which each of the firms accounts for 7%, even although structure is quite different in the two industries.

(ii) The Lorenz measure and Gini coefficient are both computed from the Lorenz curve which may be alternatively expressed as the cumulative proportional concentration curve.[1]

Figure A.6.3.1. The Lorenz curve.



Thus, in the figure, point P shows that the smallest x% of firms in the industry account for y% of total employment, in other words, the complement of the proportional concentration ratio is measured along the vertical axis.

The Lorenz measure (L) is given by the ratio of the shaded area to the triangle ABC. (For the sake of brevity and because of the close similarity of the Gini coefficient to L[2] the following discussion will relate only to the latter.)

In that it takes into account the entire size distribution, L is based on more information than is CR_5. On the other hand, it is often argued that

1. For a discussion of these measures, see Scherer, op.cit., p.50 onwards and Aitchison and Brown, op.cit., pp. 111 - 115.

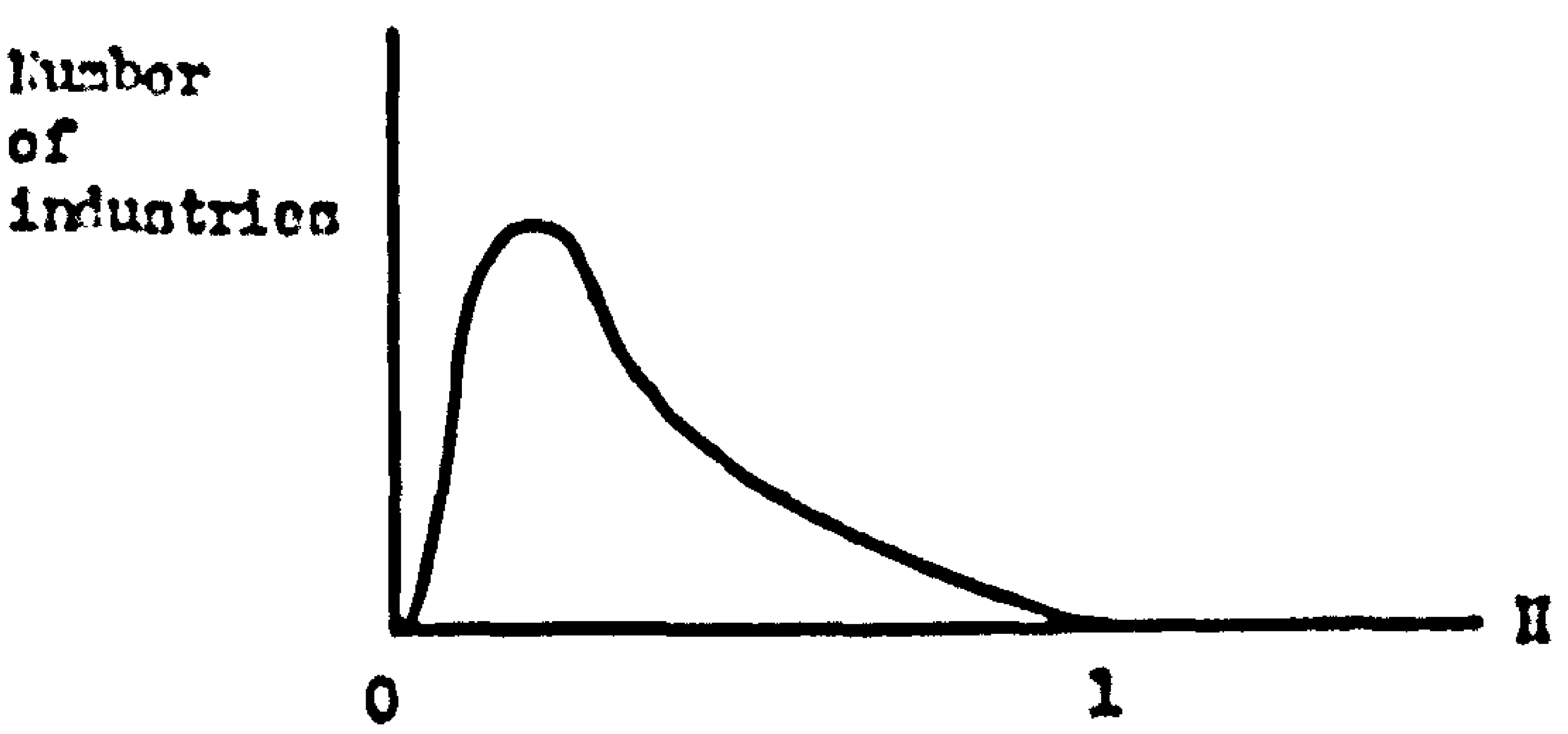
2. G = 2 S.L. where S is the arithmetic mean firm size, ibid. p.113.

it places too much emphasis on inequality, at the expense of the number of firms. For instance, an industry of 5 equally sized firms - an obvious oligopoly - would produce the same value for L (= 0) as one of 1000 equally sized firms. The upper limit for L is 1, which indicates a perfect monopoly, the lower limit is 0, which indicates perfect equality (which, in view of the above example, should not be confused with perfect competition).

(iii) It is often claimed that the Herfindahl index (H)[1] achieves the best of both worlds. It is computed as the sum of the squared employment shares for all of the firms in the industry. The extreme values are 1 (for a monopoly) and 1/n for an industry of n equally sized firms; the more inequality and the fewer the number of firms in the industry, the higher will be H. One drawback sometimes noted[2] is that the observed values for the H index tend to be distributed across industries in a rather lopsided manner: the vast majority tend to be bunched together with very low values whilst only a few industries record relatively high values.

In practice, therefore, the consequently low var (H_j) across industries may make this index a poor empirical device for differentiating between the structure of industries.[3]

Figure A6.3.2. The distribution of industries according to the H index.



1. Again, see Scherer (op.cit.) p.51 for a discussion of this measure.

2. ibid, p.52.

3. This argument is completely empty if one believes that the degree of competition is distributed across industries with a positive skew. i.e. the majority of industries being 'roughly competitive' and only a minority exhibiting oligopolistic or monopolistic tendencies. In that case, H is an ideal measure. However, there seems to be no strong reason for such a belief.

(iv) A fourth alternative, popular in recent years, is Entropy (E), a measure borrowed from Information theory.[1] This is computed as the sum of each firm's employment share weighted by the logarithm of the reciprocal of that share; thus it is not dissimilar in concept to the H index. The relationship to concentration is inverse in this case however: n equally sized firms yields a value of log n and for a monopoly, E = 0. As for H, therefore, both the number of firms and size inequalities are reflected in this measure.

(v) Finally the variance of the logarithm of firms' size within an industry has been suggested as a simple measure of concentration.[2] This has the same advantages and limitations as the Lorenz measure - it reflects the entire size distribution but totally ignores the importance of the number of firms. Again, therefore, an industry of 5 equally sized firms can not be differentiated from one of 1000 equally sized firms.

Table A6.3.1. Statistical definitions of concentration measures.

Measure of concentration.	Mathematical computation required.
1. 5 firm concentration ratio (CR_5)	$\left(\sum_{i=1}^{5} S_i\right)/n\bar{S}$
2. Lorenz measure (L)	Generally, graphical measures required to calculate L.
3. Herfindahl index (H)	$\sum_{i=1}^{n}\left(\frac{S_i}{n\bar{S}}\right)^2$
4. Entropy (E)	$\sum_{i=1}^{n}\left(S_i/n\bar{S}\right)\log\left(\frac{n\bar{S}}{S_i}\right)$
5. Variance of log size (σ_s^2)	$\frac{1}{n}\sum_{i=1}^{n}\left(\log S_i - \overline{\log S}\right)^2$

N.B. where firms are ranked in descending order of size from 1 to n, S_i = size of the firm i, $\bar{S}$ = arithmetic mean size, and $\overline{\log S}$ = the arithmetic mean of log. size.

1. See H. Theil, 'Economics and information theory,' North Holland, Amsterdam (1967) pp.290 - 3; P.Hart, 'Entropy and other measures of concentration,' Journal of the Royal Statistical Society, 1971.

2. In recent years by Hart and Prais, (1956), op.cit., and, formerly, by Gibrat, op.cit.

4. A useful simplification.

The close similarity between these five measures, which is fairly obvious on an intuitive level, may be formalised if one is prepared to assume that firm size is lognormally distributed within industries. The results of the previous appendix suggest that this is a reasonable assumption for the sample industries. At any event, small divergences from the lognormal distribution are unlikely to destroy the main thrust of the following argument.

As Table A6.4.1. shows, under these conditions, only a knowledge of n and σ_s^2 is required to calculate any of the five measures. Not surprisingly, these two parameters reflect the two aspects of concentration, mentioned as

Table A6.4.1. The definitions assuming lognormality.

Measure of concentration.	Assuming lognormal size distribution, computed as:
CR_5 [a]	$z(CR_5) = \sigma_s + z(5/n)$
L [b]	$2N(\sigma_s/\sqrt{2} \mid 0, 1) - 1$
H [c]	$e^{\sigma_s^2}/n$
E [c]	$\log n - \sigma_s^2/2$
σ_s^2	σ_s^2

a. $Z(CR_5)$ is the normit of the concentration ratio, e.g. if $CR_5 = .95$, $z(CR) = 1.645$. A proof of this result is provided at the end of this appendix.

b. A proof of this is given by Aitchison and Brown (op.cit.) p.112.

c. Proofs in Hart (1971) op.cit.

important in section 2.

Clearly, the choice between these alternatives depends on the exact relationship desired between concentration (C) and the number of firms (n) and the inequalities in their sizes (as reflected in σ_s^2.)

This may be seen more easily with the aid of two diagrams. Figure A6.4.1.

shows the relationship between each measure and σ_s^2, holding n constant, at a typical value of 100.[1] As can be seen, all five show the desirable property of $\partial C/\partial \sigma_s^2 > 0$.[2] That is, concentration increases with increasing size inequalities. However, the exact shape of the relationship is very different between the alternatives. For L and CR_5, $\partial^2 C/\partial \sigma_s^{22} < 0$: the rate of increase in concentration slows down as σ_s^2 increases. For E and σ_s^2 (of course), there are 'constant returns': an increase in size inequality has the same effect whatever the original level of σ_s^2. Only for H does concentration accelerate as σ_s^2 increases.

Tentatively, this might suggest H as the most appropriate measure. It seems most likely that small increases in size inequality will have less effect on firms' ability to collude in relatively unconcentrated industries than in ones that are already highly concentrated. For instance, a merger between large or medium-large firms in a concentrated industry may be much more effective, in this sense, than if the industry concerned is unconcentrated in the first place. Nevertheless, a reasonably open mind in this connection is probably most justified.

Figure A.6.4.2. shows the relationship between each measure and n, this time holding σ_s^2 constant, at a typical value of 2.56. In this case, only H, E and CR_5 exhibit the desirable condition: $\partial C/\partial n < 0$. Undoubtedly, one should require that increases in the number of firms reduce the level of concentration. Thus L and σ_s^2 can be ruled out as inappropriate: both are insensitive to the level of n. The three remaining measures each imply that increases in n will have a diminishing negative influence on concentration. This seems reasonable: for instance, whilst firms in a population of 20 should find it easier to collude than in a population of 100, it is unlikely that collusion will be significantly easier to achieve where there are 100

1. This is approximately the average value for n for the sample industries.
2. Recalling that E is an inverse measure of concentration.

Figure A.6.4.1. The influence of σ_a^2 on concentration (for n = 100.)

(a)

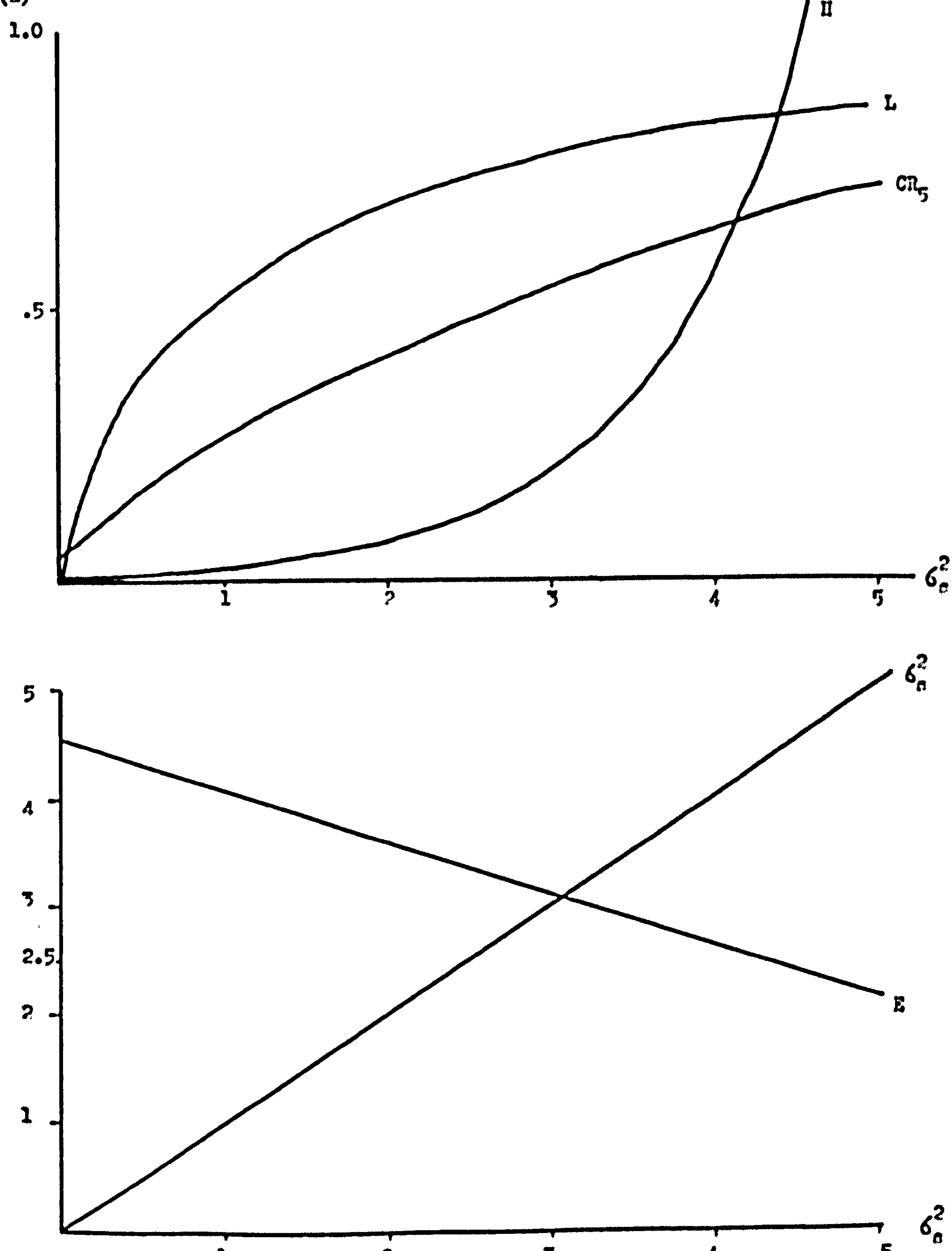
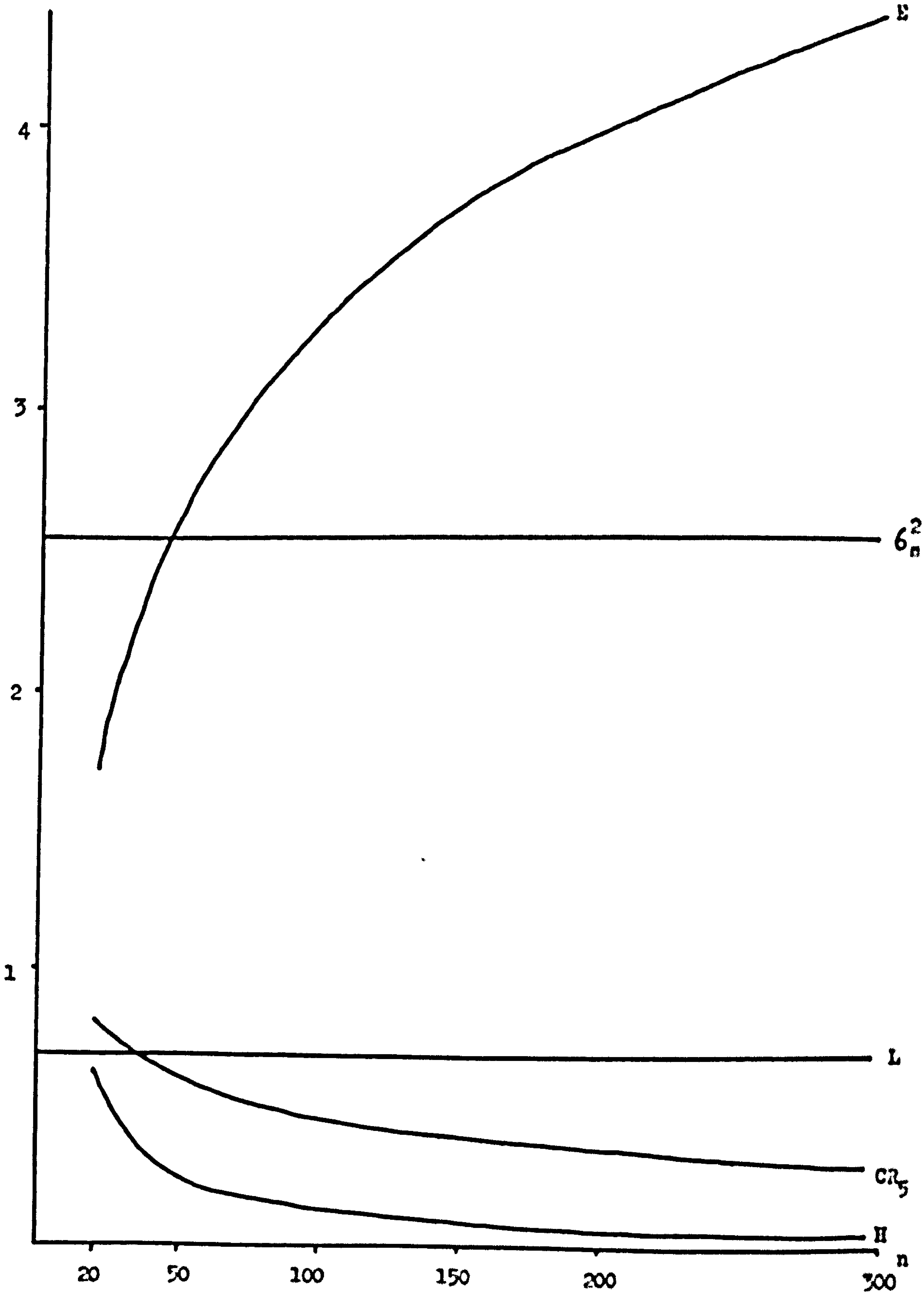


Figure A.6.4.2. The influence of n on concentration (for $\sigma_n^2 = 2.56$.)



as opposed to 200 firms (all for constant σ_n^2.) Again, the H index might be marginally preferable, in that it is least sensitive to increase in n above 100 firms. (Entropy, on the other hand, continues to increase noticeably with increases in n, even at very high levels.)

In summary, then, the H index seems to be most appropriate, but the grounds for rejecting E and CR_5 are not really conclusive. Moreover, the H index does have the disadvantage mentioned earlier which can now be confirmed by inspection of the two diagrams. For values of $\sigma_n^2 < 3$, $H < .2$ (assuming n = 100) and for values of $n > 50$, $H < .258$ (assuming $\sigma_n^2 = 2.56$.) As most industries, in practice, will probably have values for σ_n^2 and n within these ranges,[1] this does mean that the typical value for H will be very low relative to the upper limit of 1. As noted already, this implies that most industries exhibit modest concentration, with only a few showing oligopolistic or monopolistic tendencies. Furthermore, the explanatory power of H in regression analysis may be largely determined by the performance of the few industries with high H values.

5. The U index: a generalised measure of concentration.

In order to overcome these difficulties, a new measure of concentration is needed. Fortunately, only a very simple extension of the Herfindahl index is necessary.

The H index may be thought of as a special case[2] of the more general index (defined for convenience by U):

$$U = e^{a\sigma_n^2} n^{b} \qquad (A.6.5.1)$$

where a = 1, b = -1.

Another special case, which would be similar to H, but which would produce a less skewed distribution of values across industries, is a = .5, b = -.5. Alternatively, the logged form of this index may be seen as the general form of Entropy[3] and σ_n^2:

1. Within the sample, only 2 of the 13 industries have $\sigma_n^2 > 3$ and only 4 have less than 50 firms.

2. Still assuming size to be lognormally distributed within industries.

$$\log U = a\,\sigma_s^2 + b \log n \qquad (A.6.5.2.)$$

Where $E = \log U$, $a = -.5$, $b = 1$. Similarly, $\sigma_s^2 = \log U$ with $a = 1$, $b = 0$. Viewed in this way, there seems to be nothing to be gained by selecting *any* particular special form of U, or log U, if the general forms may be used instead. On the other hand, by allowing a and b to take on whatever values best fit the data concerned, the extra flexibility possible has at least two advantages. First, there is no need to employ a measure which *assumes* (albeit implicitly), that concentration is distributed across industries in any particular way. Second, the 4 partial derivatives may take on positive or negative signs.[1]

Computationally, the U index (or log U) requires only that two explanatory variables be used in the regressions in place of one. A number of arguments might be raised against such a procedure, but none of them seem to be substantive: *1) a and b are unknown.* This does not rule out the empirical use of the index, however. For example, a simple cross industry regression of profit rates against concentration:

$$\pi_j = \alpha C_j^{\beta} \qquad (A6.5.2.)$$

may be computed as $\pi_j = \alpha e^{\beta a \sigma^2} n^{\beta b}$, $\qquad$ (A6.5.3.)

where a and b take whatever values best fit the data. Certainly it would be impossible to estimate a and b, or β, from such a regression, but it is difficult to argue that estimates of a and b provide *less* information than merely $\hat{\beta}$ (where C is measured, say, by the H index.) Rarely in empirical application does β have any specific economic meaning. Indeed a regression based on (A6.5.3) would provide a test of the appropriateness of the H index for which $\widehat{(a\beta)}/\widehat{(b\beta)} = -1$.

1. That is, the first and second order partial derivatives with respect to n and σ_s^2. Having said this, of course, it is only the second order derivative for σ_s^2 which can not be confidently predicted from a priori reasoning (see above.)

2) it requires more data than the conventional measures. This may be discounted for E, H, σ_s^2 and L, but it must be acknowledged that often the only data available, in practice, for empirical size distributions is CR_5 and that individual firms' sizes are not known. Under these circumstances, however, no other measure of concentration may be used either. Moreover, if one is prepared to assume a lognormal size distribution then σ_s may be estimated from CR_5 [1] and the U index may be used. At any event, in this study, data is available for both σ_s and n for all j.

3) its use may be marred by multi-collinearity. There seem to be no theoretical reasons for n to be collinear with σ_s^2 and, indeed, for the sample industries, no significant correlation exists.

4) firm size is not always lognormally distributed. Whilst this index is only a generalised form of H and E, if size is lognormally distributed, it is still a meaningful measure of concentration even in the widespread absence of such a distribution. As has been noted already, its two constituent parts, σ_s^2 and n, may each be justified on purely economic grounds.

In summary, then, there seem to be no strong reasons for not measuring concentration by U or log U, with a and b free to vary in value.

6. Barriers to entry and product differentiation.

Far less discussion has accompanied past measurement of either barriers to entry or product differentiation. This is hardly surprising given the lack of ready appropriate data in Census reports etc. More often than not, crude dummy variables have been used which require subjective assessments. Where entry barriers have been measured by continuous variables,[2] the most

1. See table A6.4.1, row 1.

2. For a summary, see L.W.Weiss, 'Quantitative studies of industrial organization', in M.D. Intriligator, (ed.), 'Frontiers of Quantitative Economics,' North Holland (1971) Chapter 9.

popular choices have been advertising-sales ratios, capital requirements, and the extent of the scale barrier (measured, for instance, by the average size of the plants accounting for half of employment, divided by total employment.[1]) In practice, these measures either require data which is not available at the level of disaggregation required here or are only partially successful in representing the desired concept. For instance, the above measure of scale barriers seems more appropriate as a measure of concentration.

Consequently, as a very rough approximation, the median firm size in the industry[2] will be used as a catch-all measure of entry barriers. Comparing across industries, one might expect that where the absolute cost, capital requirements and scale barriers are high, so, too, will be median firm size. This measures fails, however, to represent the product differentiation barrier; unfortunately, data on advertising or any other reflection of differentiation (such as the number of products offered by each industry) is unattainable and, as is often the case, this latter aspect of structure remains unrepresented in the empirical analysis.

Thus, the index of overall competition is defined as:

$$CC_j = A\, e^{a\sigma^2_{sj}} \quad N_j^{\,b} \quad e^{c\mu_{sj}} \qquad (A6.6.1)$$

where A and c are constants and μ_{sj} is the mean of log size in the jth industry. That is, the level of competition in the jth industry is determined by the extent of concentration and of entry barriers, defined in this simple way. The expression is multiplicative in order that the effects of entry barriers are partly determined by the extent of concentration and vice-versa.[3] Had data been available from which to derive some measure

1. W. Comanor and T. Wilson, 'Advertising, market structure and performance,' Review of Economics and Statistics, vol. 49, 1967.
2. Defined by the geometric mean, which is e^{μ_s} if firm size is lognormally distributed. This choice is purely pragmatic; the arithmetic mean, given by $e^{\mu_s + \frac{1}{2}\sigma^2_s}$ in the lognormal case, could well be collinear across industries with σ^2_s, which is included in the above U index.
3. Which seems more reasonable than an additive expression for most empirical applications, e.g. the explanation of profit rates (see Weiss, op.cit.)

of product differentiation, (PD) this might have been extended to:

$$CC_j = A\, e^{a\sigma^2_{sj}}\, N_j^{\,b}\, e^{c\mu_{sj}}\, PD_j^{\,d} \qquad (A6.6.2)$$

As it stands, the index given by (A6.6.1) can provide only a partial measure of competition, but it does have two attractive features. First, data on μ_{sj}, σ^2_{sj} and N_j are generally available for most industries in practice and, second, each of these parameters of the size distribution appear already in the theoretical model. Given the shortage of degrees of freedom in the empirics of chapter 8, it is particularly convenient that CC_j can be included in the analysis without introducing new variables. More positively, it is doubtful whether more appropriate variables are available without extensive data collection at the firm-level.

Note. The concentration ratio with a lognormal size distribution (see table A6.4.1.)

If firm size is lognormally distributed, $\Lambda(\mu_s, \sigma^2_s)$ then the first moment distribution function[1] is $\Lambda(\mu_s + \sigma^2_s, \sigma^2_s)$. The former shows the number of firms at each firm size, the latter, the employment accounted for by each firm size.

Defining S_6 as the size of the sixth largest firm, then

$$CR_5 = 1 - \Lambda(S_6 \mid \mu_s + \sigma^2_s, \sigma^2_s) \qquad (A6.N.1)$$

where $\Lambda(S_6 \mid \mu_s + \sigma^2_s, \sigma^2_s)$ denotes the share of employment accounted for by all firms other than the top five.

Similarly, $1 - 5/n$ firms have size equal to or less than S_6 :

$$1 - 5/n = \Lambda(S_6 \mid \mu_s, \sigma^2_s)$$

Therefore $5/n = 1 - \Lambda(S_6 \mid \mu_s, \sigma^2_s)$ (A6.N.2)

Re-expressing (A6.N.1 and 2) using the standard normal distribution,

$$CR_5 = 1 - N\left(\frac{\log S_6 - \mu_s}{\sigma_s} - \sigma_s \mid 0, 1\right) = N\left(\sigma_s - \frac{(\log S_6 - \mu_s)}{\sigma_s} \mid 0, 1\right)$$

1. Following Aitchison and Brown's notation, (op.cit.) pp.111 - 113.

$$5/n = 1-N\left(\frac{\log B_6 - \mu_n}{\sigma_s} \;\middle|\; 0,\ 1\right) = N\left(\frac{-(\log B_6 - \mu_s)}{\sigma_s} \;\middle|\; 0,\ 1\right)$$

If $z(CR_5)$ is the normit (normal equivalent deviate) of CR_5 and likewise $z(5/n)$ for $5/n$,

$$z\,(CR) = \sigma_s - \frac{(\log B_6 - \mu_s)}{\sigma_s}$$

and $$z\,(5/n) = -\frac{(\log B_6 - \mu_s)}{\sigma_s}$$

Therefore $Z(CR) = \sigma_s + z(5/n)$

As an example, an industry of 100 firms and σ_s of 1 has a concentration ratio of 26% : $Z(CR) = 1 - 1.645 = -.645$.

Therefore CR = .26 (from normal distribution tables.)

REFERENCES

Abramowitz, M.A. "Resources and output trends in the United States since 1870", American Economic Association Papers and Proceedings, vol. 46, May 1956.

Adams, W. and Dirlam, J. "Big steel invention and innovation", Quarterly Journal of Economics, May 1966.

Aitchison, J. and Brown, J. The lognormal distribution Cambridge University Press, 1957.

Allen, T. and Reilly, V. "Getting the word around," Report of a pilot study on technology transfer to Irish industry. Alfred Sloan School of Management, Working paper 650 - 73.

Arrow, K.J. "The economic implications of learning by doing," Review of Economic Studies, June 1962.

Bailey, N. Mathematical theory of epidemics, Griffin, London, 1957.

Bain, A. The growth of T.V. ownership in the U.K. since the war. Cambridge University Press, 1964.

Bain, J. Barriers to new competition. Harvard University Press, Cambridge, (Mass.) 1956.

Bartholomew, D.J. Stochastic models for social processes, Wiley, London, 1970.

Berksen, J. "Estimate of the integrated normal curve by minimum normit chi-square," Journal of the American Statistical Association, June 1955.

Berksen, J. "Tables for use in estimating the normal distribution function by normit analysis", Biometrica, 1957.

Blaug, M. "A survey of the theory of process innovations," Economica, February 1963.

Bonus, H. "Quasi-Engel curves, diffusion and the ownership of major consumer durables," Journal of Political Economy, May/June 1973.

Bruni, L. "Internal economies of scale with a given technique," Journal of Industrial Economics, June 1964.

Cable, J.R., Chapter 4 of The U.K. economy, ed. A. Prest and C.Coppock, Weidenfeld and Nicolson, London, 1970.

Carter, C. and Williams, B. Industry and technical progress, Oxford University Press, 1957.

Carter, C. and Williams, B. Investment in innovation, Oxford University Press, 1958.

Chamberlin, E.H. The theory of monopolistic competition, Harvard University Press, Cambridge (Mass.), 7th edition, 1956.

Chow, G. "Technological change and the demand for computers," American Economic Review, 1967.

Comanor, W. and Wilson, T. "Advertising, market structure and performance," Review of Economics and Statistics, 1967.

Courant, R. Differential and integral calculus, Interscience, New York, 1937.

Cramer, J.S. Empirical econometrics. North Holland, Amsterdam, 1969.

Cyert, R. and March, J. A behavioural theory of the firm. Prentice Hall, New Jersey, 1963.

Czepiel, J. "Word of mouth processes in the diffusion of a major technological innovation," Journal of Marketing Research, 1974.

David, P.A. Technical choice, innovation and economic growth, Cambridge University Press, 1975.

Denison, E. Why growth rates differ. Brookings Institute, Washington, 1967.

Eels, F., Hazlewood, A., Knowles, K., Winsten, C. "Innovation and automation", Bulletin of the Oxford Institute of Technology and Statistics, Vol. 21, 1959.

Engewall, L. Models of Industrial Structure. Lexington Books, New York 1974.

Enos, J.L. "A measure of the rate of technological progress in the petroleum refining industry," Journal of Industrial Economics, June 1959.

Fabricant, S. Basic facts on productivity changes, N.B.E.R., Columbia University Press, New York, 1959.

Finney, D. Probit analysis, a statistical treatment of the sigmoid response curve. Cambridge University Press, 1947.

Freeman, C. The economics of industrial innovation, Penguin, London, 1975.

Gibrat, R. Les inégalités économiques, Paris, 1931.

Glaister, S. "Advertising policy and returns to scale in markets where information is passed between individuals," Economica, May 1974.

Globbermann, S. "Technological diffusion in the Canadian tool and die industry," York University Working Paper.

Gold, B., Pierce, W., Rossegger, G. "Diffusion of major technological innovations in U.S. Iron and Steel," Journal of Industrial Economics, July 1970.

Griliches, Z. "Hybrid corn: an exploration in the economics of technological change," Econometrica, October 1957.

Hakonsen, S. "Special presses in papermaking," chapter in Nabseth and Ray (1974)

Haldi, J. and Whitcomb, D. "Economies of scale in industrial plants," Journal of Political Economy, vol. 75, 1967.

Hall, R. and Hitch, C. "Price theory and business behaviour", Oxford Economic Papers, 1939.

Hart, P.E. "The size and growth of firms," Economica, 1962.

Hart, P.E. "Entropy and other measures of concentration," Journal of the Royal Statistical Society, 1971.

Hart, P.E. and Prais, S. "The analysis of business concentration: a statistical approach," Journal of the Royal Statistical Society, 1956.

Hartley, K. "The learning curve and its application to the Aircraft industry," Journal of Industrial Economics, March 1965.

Hirsch, W.Z. "Firm progress ratios," Econometrica, April 1956.

Isard, W. and Schooler, E. Location factors in the petro-chemicals industry, U.S. Department of Commerce, 1955.

Jacquemin, A. and Thisse, J. "Strategy of the firm and market structure," in Market structure and corporate behaviour. ed. K.Cowling, Gray-Mills, London 1972.

Johnson, N. and Kotz, S. Continuous univariate distributions-2. Houghton Mifflin, Boston, 1970.

Johnston, J. Econometric methods. McGraw-Hill, New York, 1963, 1st edition.

Kennedy, C. and Thirlwall, A.P. "Technical progress: a survey," Economic Journal, vol. 82, March 1972.

Leckie, A. and Morris, A. "The effect of plant and works scale on costs in the Iron and Steel Industry," Journal of the Iron and Steel Institute, May 1968.

Lilliefors, H.W. "On the Kolmogorov-Smirnoff test for goodness of fit," Journal of the American Statistical Association, 1967.

Lund,P. and Miner,D. "The investment behaviour of small firms," in Research report 11 to the Boulton Committee of Inquiry into small firms.

McGee,J.S. "*In defence of industrial concentration*," New York, 1971.

Machlup,F. "Theories of the firm; marginalist, behavioural and managerial," American Economic Review, March 1967.

Mansfield,E. "Technical change and the rate of imitation," Econometrica, 1961.

Mansfield,E. "Entry, Gibrat's law, innovation and the growth of firms," American Economic Review, December 1962.

Mansfield,E. "The speed of response of firms to new techniques," Quarterly Journal of Economics, May 1973.

Mansfield,E. *Industrial research and technological innovation*, New York, Norton, 1968.

Mansfield,E., Rapaport,J., Schnee,J., Wagner,S., Hamburger,M. *Research and Innovation in the modern corporation*, Norton, New York, 1972.

Martilla,J. "Word of mouth communication in the industrial adoption process," Journal of Marketing Research, May 1971.

Massey,F.J. "The Kolmogorov-Smirnoff test for goodness of fit," Journal of the American Statistical Association, 1951.

Metcalfe,J.S. "Diffusion of innovations in the Lancashire textile industry," Manchester School, June 1970.

Meyer,J. and Kuh,E. *The investment decision*. Harvard University Press, Cambridge (Mass.) 1957.

Miles,C. "*Lancashire textiles: A case study of industrial change*," Cambridge University Press, 1968.

Nabseth,L. "The diffusion of innovations in Swedish industry," chapter in Williams 1973.

Nabseth,L. and Ray, G.F.(ed.) *The diffusion of new industrial processes: An International study*, Cambridge University Press, 1974.

O.E.C.D. *Gaps in technology*, Paris, 1968.

Ozanne,U. and Churchill,G. "Adoption research: information sources in the industrial purchasing decision," Proceedings, Fall Conference, American Marketing Association, 1968.

Pavitt,K. *The conditions for success in technological innovation*, O.E.C.D.Paris,1971.

Pratten, C.F. *Economies of scale in manufacturing industry*, Cambridge University Press, 1971.

Pyatt,F.G. *Priority patterns and the demand for household durable goods*, Cambridge University Press, 1964.

Quandt,R. "On the size distribution of firms," American Economic Review, June 1966.

Ray, G.F. "The diffusion of new technology," National Institute Economic Review, May 1969.

Rogers,E. *Diffusion of innovations*. Free Press of Glencoe, New York, 1962.

Rosenberg,N. *The economics of technological change*. Penguin Readings, London 1971.

Rowley, C.F. *Antitrust and Economic efficiency*, McMillan, London,1973.

Salter,W. *Productivity and technical change*. Cambridge University Press, 1960.

Saving,T. "The four parameter lognormal, diseconomies of scale and the size distribution of manufacturing establishments," International Economic Review, January 1965.

Scherer,F. Industrial market structure and economic performance. Rand McNally, Chicago, 1971.

Schmookler,J. Invention and economic growth, Harvard University Press, Cambridge (Mass.) 1966.

Schumpeter,J. The theory of economic development, Harvard University Press, New York.

Schumpeter,J. Capitalism, socialism, and democracy, Harper and Row, New York,1942.

Schwartzmann,D. "Uncertainty and the size of firm," Economica, 1963.

Silbermann,I. "On lognormality as a summary measure of concentration," American Economic Review, September 1967.

Simon,H. and Bonini,C. "The size distribution of business firms," American Economic Review, September 1958.

Smith,R.J. "Shuttleless looms," chapter in Nabseth and Ray (1974)

Solow,R.M. "Technical change and the aggregate production function," Review of Economics and Statistics, vol.39, August 1957.

Stigler,G. "The kinky oligopoly demand curve and rigid prices," Journal of Political Economy, October 1947.

Stoneman,P. "On the change in techniques - a study of the spread of computer usage in the U.K., 1954-70," Ph.D. Thesis, Cambridge, 1974.

Stoneman,P. "The choice of technique: the example of computerisation," Warwick Economic Research Paper no. 46, 1974.

Swann,P. "The international diffusion of an innovation," Journal of Industrial Economics, September 1973.

Taylor,W.F. "Distance functions and the regular best asymptotically normal estimates," Annals of Mathematical Statistics, 1953.

Theil,H. Economics and information theory. North Holland, Amsterdam, 1967.

Webster,F. "Informal communication in industrial markets," Journal of Marketing Research, May 1970.

Weiss,L. "Quantitative studies of industrial organisation" in M.D. Intriligator(ed.) Frontiers of quantitative economics, North Holland, Amsterdam,1971

Wigley,K. The demand for fuel 1948 - 75, Vol. 8, A programme for growth. Cambridge University Press, 1968.

Williams,B.R.(ed.) Science and technology in economic growth. MacMillan, London 1973.

Williamson,O. "A dynamic theory of inter-firm behaviour," Quarterly Journal of Economics, May 1965.

See also Appendix one, pp. 53 - 60 for a list of technical and scientific references.